

# Poverty and civil war

*An assessment of four prominent explanations of  
the per capita income-civil war relationship*

**Helge Holtermann**



Masteroppgave

Institutt for statsvitenskap

UNIVERSITETET I OSLO

November 2007

## Abstract

Civil war is highly concentrated in poor countries. Nearly all recent quantitative studies find a strong relationship between per capita income and risk of civil war onset, controlling for other explanatory variables. How can we explain this? Four different explanations have reached prominence in the quantitative literature. First, the *grievance explanation* holds that people in poor countries for various reasons are likely to be more frustrated with their government, making them more willing to rebel. Second, the *state weakness explanation* holds that poor countries are more prone to civil war because their state organizations are less capable of controlling social relationships in general and defeating rising insurgencies in particular. Third, the *greed explanation* holds that poor countries have a higher risk of civil war because lower economic opportunities imply a greater incentive for people to become rebels. Finally, what I call the *ecological explanation* holds that civil wars more often occur in poor countries because they have social and physical environments more conducive to guerilla warfare - the usual military tactic of 20<sup>th</sup> century insurgents.

Previous studies have largely focused on only one of these explanations, and often not done rivaling explanations justice when interpreting empirical evidence. This study is the first to systematically assess their relative strength through a theoretical and empirical analysis. In the theoretical part I discuss plausible lower-level arguments related to each explanation. By specifying the causal structure of these arguments I derive several hypotheses which help to distinguish between the explanations. These are tested using logistic regression analysis of quantitative data for 161 countries from 1945 to 1999.

My main conclusion is that the ecological explanation appears to be the strongest of the four explanations. Two of the ecological variables - the proportion of people living in urban areas and the extensiveness of telecommunication infrastructure - can empirically account for the entire per capita income effect on the risk of civil war. The two most prominent explanations in recent quantitative literature, state weakness and greed, find surprisingly little empirical support. Only one of the grievance hypotheses was to some extent supported by the evidence.

# Table of contents

Abstract.....	1
Table of contents .....	2
List of figures and tables .....	3
Preface .....	4
Chapter 1 Introduction .....	5
1.1 The puzzling poverty-civil war relationship .....	5
1.2 Research gap and research question .....	6
1.3 Research design and structure of the study .....	7
Chapter 2 Definitions of some central concepts .....	12
2.1 The State.....	12
2.2 State capacity and strength.....	13
2.3 Civil war.....	14
Chapter 3 Theoretical analysis .....	16
3.1 The grievance explanation .....	16
3.2 The state weakness explanation .....	23
3.3 The ecological explanation.....	41
3.4 The greed explanation .....	46
3.5 Summary of hypotheses .....	53
Chapter 4 Statistical method .....	56
4.1 Logit analysis of cross-national time-series data .....	56
4.2 The multicollinearity problem.....	57
4.3 The logic and structure of the empirical tests .....	59
4.4 Methodological problems and remedies .....	61
4.5 Choosing a set of control variables .....	64
Chapter 5 Indicators and data.....	66
5.1 Civil war onset .....	66
5.2 Income per capita .....	70
5.2 Grievance indicators.....	71
5.3 State capacity indicators.....	73
5.5 “Ecology” indicators .....	75
5.6 Economic opportunity indicators (greed explanation) .....	76
5.7 Control variables .....	77
Chapter 6 Empirical analysis.....	79
6.1 From the Fearon-Laitin model to a new basic model.....	79
6.2 Testing the grievance hypotheses.....	81
6.3 Testing the state weakness hypotheses.....	91
6.4 Testing the ecological hypotheses.....	100
6.5 Testing the greed hypotheses .....	104
Chapter 7 Conclusion .....	111
7.1 Summary of findings.....	111
7.2 Discussion .....	116
References .....	118
Appendices .....	122
Appendix 1. Descriptive statistics.....	126
Appendix 2. Multicollinearity assessment .....	127

# List of figures and tables

## Figures

<i>Figure 1.</i> Causal diagram showing the grievance argument that increased extraction in states with initially low extraction levels explains the poverty-civil war relationship .....	21
<i>Figure 2.</i> Causal diagram showing the argument that increased extraction in poor countries explains the poverty-civil war relationship .....	22
<i>Figure 3.</i> Causal diagram showing the argument that the causal effect of income per capita on risk of civil war goes through public goods provision .....	23
<i>Figure 4.</i> Causal diagram showing the argument that state capacities explain the poverty-civil war relationship .....	30
<i>Figure 5.</i> Causal diagram showing the argument that decrease of revenues in countries with weak states explains the poverty-civil war relationship.....	37
<i>Figure 6.</i> Causal diagram showing the argument that state building in weak states explains the poverty-civil war relationship .....	41
<i>Figure 7.</i> Causal diagram showing the argument that ecological variables explain the poverty-civil war relationship .....	45
<i>Figure 8.</i> Causal diagram showing the greed / opportunity cost model .....	53
<i>Figure 9.</i> Graph showing the number of civil war onsets and ongoing civil wars by year from 1945 to 1999 .....	69
<i>Figure 10.</i> Estimated risk of civil war onset (in log-odds) by tax ratio change and tax ratio, based on Model 3-3 .....	85
<i>Figure 11.</i> Estimated risk of civil war onset (in log-odds) by tax ratio change and income per capita, based on Model 3-2 .....	86

## Tables

<i>Table 1.</i> Overview of hypotheses and corresponding explanations .....	54
<i>Table 2.</i> Replication and changes to Fearon and Latin's (2003a) logit model of determinants of civil war onset .....	80
<i>Table 3.</i> Test of H1 and H2: Estimating the conditional effect of change in tax ratio on the risk of civil war onset .....	84
<i>Table 4.</i> Test of H3: Estimating the effect of primary schooling on the risk of civil war onset .....	90
<i>Table 5.</i> Test of H4.1 and H4.2: Estimating the effect of primary schooling on the risk of civil war onset .....	93
<i>Table 6.</i> Test of H4.3 and H5: Estimating the effect of government share and decrease of government share on the risk of civil war onset.....	96
<i>Table 7.</i> Test of H4.4 and H8.5: Estimating the effect of tax ratio on the risk of civil war onset .....	99
<i>Table 8.</i> Test of H6 and H7: Estimating the effect of settlement structure and telecommunication infrastructure on the risk of civil war onset.....	101
<i>Table 9.</i> Test of H8.2, H8.3, and H8.4: Estimating the effect of economic growth, male secondary schooling, and income inequality on the risk of civil war onset.....	106
<i>Table 10.</i> Test of H8.1: Estimating the effect of per capita income on the risk of civil war onset .....	109

## Preface

The thesis was written at the international Peace Research Institute, Oslo (PRIO). I am grateful to PRIO for providing a pleasant and very stimulating environment. Several relevant brownbag paper presentations there have made lunch an inspiring event. Particularly helpful were the comments I received when presenting an early draft of my thesis at the Center of the Study of Civil War (CSCW) at PRIO.

I want to give special thanks to my supervisor, Håvard Hegre, who has shown great involvement in my project from the very beginning, and given me indispensable advice. I also want to thank PRIO's excellent librarian, Odvar Leine, for quickly providing me with all the books I needed. Thanks also to Joachim Carlsen for helping me out with statistical software, and to Scott Gates, Ragnhild Nordås, and Hanne Fjelde for useful comments and advice. I am also grateful to PRIO for the scholarship that allowed me to work with the thesis full-time. Finally, I want to thank people important in my non-academic life, and particularly Kristine, for her constant moral support.

Remaining errors are my responsibility alone.

Number of words (all included): 46 671

# 1 Introduction

## 1.1 The puzzling poverty-civil war relationship

Civil war is “the problem of the poor”, according to Sambanis (2002:216). Poor countries have over time been much more likely to experience civil war than wealthy countries.

Virtually all recent quantitative conflict studies find a strong relationship between per capita income and risk of civil war onset, controlling for other explanatory variables (Hegre and Sambanis 2006:533).<sup>1</sup> Despite agreement on this empirical relationship, our understanding remains limited. Several researchers have cited the per capita income finding in support of their favoured theory, usually offering brief theoretical arguments and little empirical evidence. Consequently, as of today we do not know which is the strongest of at least four widely different explanations. The empirical poverty-civil war relationship thus constitutes an unsolved theoretical puzzle.<sup>2</sup>

Perhaps the most intuitive explanation derives from *grievance theory*. The simplest grievance argument holds that a life in poverty makes people more frustrated, and thus more likely to rebel. This account resonates well with a common conception of civil war - the outcome of an armed popular uprising against injustice. In the academic literature this argument finds little support, however. Relative deprivation theory emphasizes that people do not rebel merely due to want of material goods. Frustration may lead to political violence primarily when people hold the government responsible for the deprivation of something they believe is rightfully theirs (Gurr 1970). I suggest two different grievance explanations based on relative deprivation theory. First, an *increasing level of taxation* may create grievances that are particularly strong in poor countries, because it conflicts with prevailing norms. Second, *lack of public goods provision* by the state in poor countries may also be a viable grievance explanation.

A second prominent explanation is state-centered. Fearon and Laitin (2003a) interpret per capita income as a proxy for *state weakness*. Poor countries, they argue, are characterized by state organizations with low administrative, financial, military, and police capacity. This makes armed insurgency a more feasible and attractive option for dissidents or opportunists. Their interpretation is thus more Hobbesian than economic: “Where states are relatively weak

---

<sup>1</sup> We use per capita income and GDP per capita interchangeably.

<sup>2</sup> By poverty we mean a low national income level compared to other countries, as usual in the quantitative literature.

and capricious, both fears and opportunities encourage the rise of would-be rulers who supply a rough local justice while arrogating the power to “tax” for themselves and, often, for a larger cause” (Fearon and Laitin 2003a:76).

A third explanation similarly holds that poverty *per se* does not make low-income countries more prone to civil war. Kocher (2004) argues that per capita income is a proxy for *settlement patterns*. Poor countries tend to be less urbanized than rich countries. This may explain the poverty-civil war relationship because *insurgency* is hardly possible in urban areas. Insurgency is here understood as a strategy based on guerrilla warfare and local political control. Urban environments are unfavourable to guerrillas because they do not facilitate hiding. Obtaining political control of large cities is also unfeasible for most insurgent organizations. Poor countries may thus be more prone to civil war because of the opportunities for insurgency offered by their characteristic environments. Following Kocher, I call this an *ecological explanation*.<sup>3</sup>

The fourth explanation I consider derives from the so-called *greed theory*. I find that of these four explanations, only the greed explanation holds that poverty has a strong *direct* and *unconditional* causal effect on the risk of civil war. Collier and Hoeffler (1998; 2004) argue that poverty increases both the incentive and the opportunity for armed rebellion. Poor people with low opportunities in the labour market have more to gain than those well off by enlisting as rebel soldiers. This in turn creates a higher supply and lower cost of labour for rebel organizations. Thus, the greed explanation holds that poor countries are more prone to civil war because poverty makes rebellion more viable and profitable.<sup>4</sup>

## 1.2 Research gap and research question

It is difficult to a priori assess whether each of these explanations may rightfully claim per capita income as a proxy for their causal variable. To empirically distinguish between the explanations we need theoretically specific indicators for the various causal variables. There have been a few attempts to test more specific measures in large-N cross-national studies (e.g., Kocher 2004; Collier and Hoeffler 2004; Fearon and Laitin 2003b). I argue that these

---

<sup>3</sup> This name is not much used today, but I find it the most appropriate here. Ecological social theories emphasize how the physical and social environment affects human behaviour (Sprout and Sprout 1965).

studies have largely promoted their one favourite explanation, however, and done little justice to the other explanations when suggesting testable implications.<sup>5</sup> Moreover, these studies investigate the determinants of civil war onset in general - a much broader research question than mine. There has not been any quantitative study concentrating solely on the poverty-civil war relationship.

I also believe too little has been done to link the relevant theoretical literatures with the quantitative empirical studies. The theoretical arguments are often briefly presented in the quantitative studies without much discussion of underlying assumptions and logical deductions. Also, the causal structure of the arguments is usually left unspecified. This leaves us wondering whether the income per capita effect is direct, indirect, spurious, or conditional on other variables. More thorough theoretical analysis may help us disclose inconsistency in the arguments, derive novel testable implications, and sharpen the specification and interpretation of statistical models.

The grand question which this study will explore is why civil war has occurred much more frequently in poorer than in richer countries over the last few decades. I approach this by focusing on the explanations given in the quantitative literature. The research question I try to answer is *to what extent can the four different categories of explanations described above - grievance, state weakness, ecology, and greed, - account for the empirical relationship between poverty and the risk of civil war onset?* “Extent” should not be taken as an indication that I will statistically estimate the precise relative strength of the explanations. As I discuss below, some fundamental problems make this virtually impossible. I try to evaluate the overall strength and plausibility of each explanation based on a theoretical and empirical analysis.

### **1.3 Research design and structure of the study**

This research question may seem overwhelming for a master thesis. I chose to pursue it after some deliberation. Reading the quantitative literature on civil war I soon found the poverty-civil

---

<sup>4</sup> Note that greed theory contains other aspects than this argument about opportunity cost of rebellion. I only consider the *greed explanation* for the poverty-civil war relationship, not *greed theory* in its entirety.

<sup>5</sup> For example, Collier and Hoeffler (2004:576) use land and income inequality and ethnic and religious diversity as proxies for grievance. We argue that these proxies do not capture grievances well, and that income inequality should rather be a measure of their greed argument.



war puzzle interesting. Much more work is done on the relationship between democracy and civil war than on the empirically stronger relationship between per capita income and civil war. Having chosen a topic, I was not sure whether confronting the puzzle by comparing several explanations was practicable. I therefore considered focusing on one explanation. This would have shared the limitations of all the other “one-explanation” studies, however. Since it is likely that more than one of the explanations is partly correct, we should try to assess their *relative strength*. The Popperian method of falsifying incorrect theories, leaving us with the correct one, is probably not viable. I consider a Lakatosian approach more expedient. Tests should be a struggle between rivaling theories and empirical evidence, according to Lakatos (1970:115). Such an analysis of the relative strength of the explanations is clearly best done by explicit comparison in one and the same study. I thus concluded that the potential analytical gains made the troubles of taking on this larger task worthwhile.

The first part of my study is *theoretical*. In Chapter 2 I define central concepts. Chapter 3 provides a theoretical investigation, structured according to the four explanations. For each explanation, I review the literature and discuss whether it is logically consistent and rests on plausible assumptions. Then I derive some lower-level arguments. To the extent possible, I specify the causal structure of these arguments illustrated by causal diagrams. Finally, empirically testable hypotheses are derived from each argument. The last section presents an overview of the hypotheses from all the explanations.

The four explanations I consider are clearly not the only possible explanations of the poverty-civil war relationship. Theories of civil war abound, and date back at least to Thucydides and the political scholarship of ancient Greece. Consequently, I had to focus on a few explanations out of many. I found it reasonable to take the quantitative civil war literature, from which my research question is drawn, as a starting point. Theories that have already been severely weakened by studies in this literature were not included.<sup>6</sup> Neither did I consider theories that are very difficult to test in cross-national analysis, such as theories of how social capital is related to conflict. The four explanations were chosen primarily because they are prominent in the quantitative literature, and because they allow for testing.

---

<sup>6</sup> For example a theory holding that poor countries are more prone to civil war because of their ethnic diversity (cf., Fearon and Laitin 2003a:83-83 and Collier and Hoeffler 2004:581).

The quantitative conflict literature does not suffice, however, when it comes to understanding this essentially *theoretical* puzzle. Theoretical analysis constitutes a small part of most quantitative studies. I therefore decided to consult other literatures for assessing the theoretical foundation of the explanations. Particularly, I draw upon the formal economic conflict literature (e.g., Grossman 1991; Hirsleifer 2001), the literature on weak states and state-building (e.g.; Herbst 2000; Reno 1998), the classics on rebellion and revolution from which grievance theories derive (e.g., Gurr 1970; Scott 1976), and the literature on insurgent and counterinsurgent strategy in civil war (e.g.; Leites and Wolf 1970; Kalyvas 2006).

An important part of the theoretical chapter is the specification of the explanations' causal structure. Underlying my attempt to do this is the assumption that per capita income can be seen as a meaningful causal variable. Since it has been used as a proxy for so many different theoretical concepts, there is much confusion surrounding its substantive content. It is often considered merely a diffuse catch-all measure. I think it is useful to hold on to its most intuitive and straightforward meaning – an indicator of poverty or generally low income levels in society. This is a more direct and less controversial interpretation than to consider it a proxy for state military capacity, for instance. By finding more specific measures for the other concepts, such as state capacity, per capita income may expediently be kept as a measure of poverty and low economic opportunities.

Perhaps because they have conceived this differently, the proponents of the different explanations have not specified whether they see poverty as causally related to civil war. Considering an archetypal form of the explanations, I argue that only the greed explanation implies a direct and unconditional causal effect of poverty. The state capacity, grievance, and ecological explanations hold that certain variables (e.g., state capacities or urban settlement structure) somehow account for the negative empirical relationship between per capita income and the risk of civil war onset. A few of these explanatory variables are hypothesized to be *intervening*, implying that the per capita income causal effect is indirect. Other variables are thought to be *confounding*, which means that the per capita income effect is actually spurious (not causal). Finally, some variables are thought to be *interacting* with per capita income, thus making the per capita income effect conditional. Testable hypotheses can be derived from these arguments. If a variable is intervening or confounding, the effect of per capita income should be weakened when statistically controlling for it. A conditional effect can be assessed by testing an interaction term between per capita income and the conditioning variable.

The second part of the study is *empirical*. In Chapter 4 I defend and account for the statistical method and the logic of hypothesis testing. The empirical measures of the dependent variable and independent are described and discussed in Chapter 5. In Chapter 6 I successively test the hypotheses derived from the four explanations. Finally, in Chapter 7 I sum up the findings and discuss what inferences can and should be made.

I assess the explanations empirically through the tests of the corresponding hypotheses. There are some limits to which inferences can be made from these tests. I may not *precisely* estimate the explanations' relative strength in the empirical analysis for two main reasons. The first is a data problem; we lack highly valid and reliable measures for some of the theoretical variables. I can not avoid this problem which is plaguing quantitative civil war research. I test a few novel indicators that I believe are more directly related to the theoretical concepts than previously used indicators. However, the validity and reliability of these indicators can also be questioned. The second problem derives from social reality itself. Some of the measures of the different explanatory variables are highly correlated. Hence, the explanatory variables can only be distinguished to a certain extent. This is a challenge for testing some, but far from all, the hypotheses.

The method used to test the hypotheses is logistic regression analysis of cross-national time-series data. I do not extensively discuss the choice of a quantitative approach. A case-study design would be inadequate for answering my research question because it does not allow for control of several alternative explanations. As we want to make inferences concerning many variables, we need to include many observations.<sup>7</sup> Using quantitative data at a more disaggregated, regional level is probably the most interesting alternative to our approach. This offers the opportunity to assess causal mechanisms more directly, and to test the hypotheses on new and more fine-grained data. However, the datasets on a sub-national level are in their infancy. It would therefore be even harder to find data for all my independent variables taking such an approach.

My main conclusion is that the ecological explanation appears to be the strongest of the four explanations. Two of the ecological variables - the proportion of people living in urban areas

---

<sup>7</sup> For a discussion of the problem of more inferences than observations, see King et al. (1994:119-121).

and the extensiveness of telecommunication infrastructure - can empirically account for the entire per capita income effect on the risk of civil war. This suggests that poor countries are more prone to civil war due to their typical social and physical environment, which makes insurgency and guerrilla warfare more feasible. The two most prominent explanations in the quantitative literature, state weakness and greed, find surprisingly little empirical support. Only one of the grievance hypotheses was to some extent supported by the evidence.

## 2 Definitions of some central concepts

In this chapter I define some central concepts that are often attributed different meanings: the state, state capacity, and civil war. The meaning of state capacity is a particularly disputed. “State failure” is sometimes used as a nearly all-encompassing concept, including the phenomenon of our interest: civil war (e.g., Esty et al. 1995:1; Rotberg 2004:1). Clearly, if civil war is a defining characteristic of failed states, it becomes meaningless to look for a causal relationship between state failure and civil war. The concept of civil war is neither unproblematic. The final section provides a definition and an argument for studying it separately from other types of collective violence.

### 2.1 The state

There are wide disagreements concerning how to define and understand the state. While historians and anthropologists often use a broad definition of the state, many political scientists and sociologists follow Max Weber and opt for a narrower definition (Chabal and Daloz 2006:226). According to Weber’s famous definition, “a state is a human community that (successfully) claims the *monopoly of the legitimate use* of physical force within a given territory” (Weber 1964:78, emphasis in original). Using Weber’s more restrictive definition is problematic for our purpose. First, Weber’s definition deprives many of the political entities that are normally called states of their statehood, as many of them have no monopoly of violence, and probably no monopoly on the legitimate use of it either (Chabal and Daloz 1999). This creates large difficulties for our study, which involves general theories of state-citizen interaction and empirical analysis of nearly the entire universe of juridical “states”. Second, Weber’s definition is unsuited for studying the relationship between state strength and civil war, since, following a strict interpretation; a “state” would cease being a state if civil war occurred, as the state would no longer hold a monopoly on (legitimate) violence.

I use a more encompassing definition of the state, including all entities that are today legally recognized as sovereign states. The state is here broadly defined as the organization which possess external sovereignty, and which *claim* the right to make and enforce the rules that govern people within a demarcated territory. Important state institutions are the executive and legislative branches of government, civil service, courts, armed forces, and police. States, according to this definition differ markedly in terms of how *successful* they are in asserting

authority over their territory and population.<sup>8</sup> However, despite such empirical differences among juridical states, pointed out by Jackson and Rosberg (1982), there are also similarities speaking in favor of a broad definition. Migdal (1988:xiii) argues that also weak post-colonial states have assumed large importance in society: “Even in the most remote corners of those societies with the newest states, the personnel, agencies, and resources of the state have reshaped political and social landscapes.” Analysis of the state’s impact on various social phenomena should thus not be confined only to the modern Western part of the World.

## 2.2 State capacity

Suggesting a definition of state capacity or strength is even more controversial than defining the state. Some argue that it is futile to compare the capacity of all states, and rather suggest a more contextually and historically fine-grained analysis (e.g., Evans et al. 1985:352-353). I acknowledge that global comparisons of state capacity face large difficulties, but think the possible gains of such an approach makes it recommendable. My research questions can hardly be answered without this. Also, I would argue that states have much in common, and their capacities to solve central challenges that they all face can and should be compared.

Huntington (1968:1) famously stated that “[t]he most important political distinction among countries concerns not their form of government but their degree of government.” I conceive of state capacity in terms of Huntington’s second dimension: the degree to which governments control and regulate society. This conception differs from approaches that see state capacity as the degree to which state leaders are autonomous from societal influences (e.g., Krasner 1978). State capacity must neither be conflated with the strength of a particular government. The Zairean state, for example, was extremely weak, although Mobutu stayed in power for 32 years. I see state capacity much like Migdal (1988:4): the ability “to *penetrate* society, *regulate* social relationships, *extract* resources, and *appropriate* or use resources in determined ways”. I also draw on Organski and Kugler’s (1980) conception, which focuses on extraction. They define state (or “political”) capacity as the “[p]enetration of the national society by central governmental elites to control as many subjects/citizens as possible within the political jurisdiction of the state; and the capability of the government to extract resources from society” (Organski and Kugler 1980:72).

---

<sup>8</sup> Weber himself probably acknowledged the difficulty of achieving total control over violence, as he writes

I understand state capacities in this relatively restrictive sense: as the degree to which the agents and institutions of the state are able to penetrate, control, and extract resources from society. The extent to which the state provides public goods such as education, welfare services, democracy, and rule of law is excluded from this definition. This contrasts with approaches that see the provision of various public goods or the adoption of developmental policies as state capacities, and make them the definitional core of the concept (e.g., Englebert 2000). Admittedly, strong states will on average provide more public goods than weak states. But it makes theoretical sense to distinguish the extent to which the state serves the citizenry from its capacity or strength. A strong state need not be a benevolent one.

## **2.4 Civil War**

There is no single authoritative definition of civil war. This is not only due to differing operationalizations of the concept. There is no clear consensus on what a civil war is, and what separates it from other types of collective violence, such as coups, terrorism, genocide, or organized crime (Sambanis 2004a:268). Tilly (2003:13) suggests that civil war has no clearly distinct ontology, and it should therefore not be studied separate from other forms of collective violence. I disagree, as most conflict researchers do (Gates 2002:3-4).

I follow Small and Singer's (1982:210) ideal-type definition of civil war as an armed conflict where (a) military action is internal to the state, (b) the national government actively participates in the struggle, and (c) both sides give effective resistance. They think the main distinction between civil war and interstate war is that fighting happens within the territory of one sovereign state. Another difference could be added: International law does not apply to civil wars. Civil war is distinguished from other forms of internal collective violence by the requirement that both parts must resist effectively (excluding pure genocides). I also add to their definition that the fighting must be sustained and killings rise above a certain threshold, in our operationalizations 100 deaths per year and 1,000 cumulative deaths over the entire war (excluding coups, riots, terrorism (usually), and organized crime)..

Civil war, as defined above, should be studied as a distinct phenomenon because it probably has several distinct causes. For a civil war to happen, first a rebel group capable of challenging the government militarily must be built up. Second, the rebels and the

---

“successfully” in parentheses.

government must fail to come to a negotiated agreement, and rather opt for war. This definition hence suggests that we should study under what circumstances a rebel group can and will organize, and rise to the level of challenging the state. We should also investigate what could make the government and the rebels choose to fight instead of negotiate a settlement.

Sambanis (2004b:816) may still be right that the definition above is “deceptively straightforward”. In reality, many armed conflicts do not neatly fit into the ideal-type definition of civil war. Some civil wars grow out of coups or riots. In other cases, it is hard to tell whether a civil war or genocide is happening, as Cambodia from 1975-1979. It may also be difficult to separate between a rebel group and an organized gang of criminals, as with the Revolutionary United Front (RUF) in Sierra Leone. Further complicating matters is the tendency of many civil wars to “spill over” to neighbouring countries. Several conflicts that are usually coded as civil wars also have significant foreign military presence (Sambanis 2004:268-271). These issues pose large problems for operationalizing and coding civil wars. I still uphold that civil war should be analyzed as a distinct phenomenon, although its relationship with other forms of collective violence should be further studied.



## 3 Theoretical analysis

This chapter discusses the theoretical foundation and empirical implications of the four explanations. Drawing on relevant theoretical literatures I assess the assumptions and logic of the four explanations and specify plausible lower-level arguments related to the explanations. I then derive a causal structure from these arguments to the extent possible, shown by causal diagrams. These provide a basis for deriving testable hypotheses. The chapter is structured in sections according to the four explanations. The final section sums up all the hypotheses and their relation to the explanations.

### 3.1 Grievance explanations

*“(...) for as long as a man thinketh himself well, and that the present government standeth not in his way to hinder his proceeding from well to better; it is impossible for him to desire the change thereof.”* Hobbes (1969:168)

The theory that people rebel primarily because of discontent, deprivation, and grievances has many adherents.<sup>9</sup> Although few scholars have stated it as bluntly, a widespread explanation of the poverty-civil war relationship holds that people living in poor countries feel more deprived, and this increases their propensity to rebel. Two fundamental questions must be answered to assess such a grievance explanation: First, why should grievances and discontent make people rebel? Second, what factors could make people in poor countries generally more aggrieved than people in wealthier countries?

#### 3.1.1. From grievances to rebellious action

In his classic, *Why Men Rebel*, Ted R. Gurr provides one answer to the first question. He suggests that the causal link from discontent to political violence goes through a bio-psychological “frustration-aggression mechanism”. This theory holds that people have a biological disposition to use violence against the agent believed to cause ones’ frustration.<sup>10</sup> Violent political action is thus not rationally chosen, but is the consequence of peoples’ uncontrollable reactions to perceived injustice: “discontent provides an innately nonrational (but widely rationalized) impetus to violence, empirically and analytically distinguishable from actors’ estimations of the utilities of violence” (Gurr 1970:326). On this theory, if people

---

<sup>9</sup> Even Hobbes, often regarded as one of the first rationalist scholars, regards discontent as the first cause of rebellion. The second is normative justification of rebellion, and the third is hope of rebel success (Hobbes 1969:168).

believe that the government or other state agents are causing them deep frustration, they will most likely turn to armed rebellion irrespective of its consequences.<sup>11</sup>

Gurr's frustration-aggression theory is not held by all scholars within what Lichbach (1989) named the "Deprived Actor" (DA) research program. A grievance theory of rebellion need not assume that frustrated people follow bio-psychological impulses in complete disregard of goals and outcomes. It may hold that people are rational and goal-oriented, but that their preferences over outcomes are influenced by grievances and norms. One way of modeling this is that individuals having grievances against the government receive a psychological satisfaction from participating in rebellion. In such a model, grievances increase the likelihood of rebellion by entering as one of several variables in a rational person's utility calculus.<sup>12</sup>

I have thus identified two possible causal mechanisms linking grievances to rebellion. Both these theories escape the Olsonian collective action critique of "public goods theories" of rebellion (cf. Tullock 1971; Olson 1965). This critique holds that rational people will not contribute to rebellion if rebels' only benefit comes from producing a public good (e.g., dethroning a bad government). As one rebel soldier's contribution to rebellion is probably negligible to the likelihood of rebel victory, and non-rebels will also benefit from the public good produced by rebellion, rational people will choose to "free-ride"; they will not take the individual costs of participating in the rebellion, but still benefit from the eventual public goods produced by it. Gurr's theory escapes this critique because it is irrationalist; aggrieved people rebel on instinct, not in order to produce a public good. The second, more rationalist grievance theory escapes the Olsonian critique because it assumes that individuals receive a *private psychological benefit* from participating in rebellion – they do not rebel only to produce a public good.

### **3.1.2 Grievances in developing countries – taxes and public goods provision**

It is commonly claimed that the scope and intensity of grievances generally increase with decreasing level of economic development. Muller et al. (1991:1262), for example, state as a fact that people feel more deprived in poor countries: "Almost all of the major insurgencies and

---

<sup>10</sup> It is based on studies in experimental psychology referred to in Gurr (1970:23).

<sup>11</sup> Gurr does not argue that frustration necessary leads to violence, or that everyone disregards the consequences of their actions when frustrated. He holds, however, that "[i]f frustrations are sufficiently prolonged or sharply felt, aggression is quite likely, if not certain to occur" (Gurr 1970:36-37).

revolutions in the second half of the twentieth century have taken place in the less-developed countries of the Third World, where objective levels of deprivation are much higher than in advanced industrialized countries.” Relative deprivation theory, however, does not point to any simple and direct relationship between poverty, deprivation, and civil violence.

Gurr (1970:24) argues that discontent and grievances derive from *relative deprivation*, defined as “actors’ perception of discrepancy between their value expectations and their value capabilities”. Value expectations are defined as the goods and conditions of life that people think they are rightfully entitled to, whereas value capabilities are the goods and conditions that they believe themselves actually capable of achieving. In Gurr’s theory, individuals form their value expectations by comparison to various *reference points*; for example the individual’s past condition, that of other persons or groups, or an abstract norm of justice. Gurr (1970:25) makes no assumptions concerning which reference point is the most salient. This makes his theory virtually non-falsifiable, and of little interest in my view. If individuals are assumed to have entirely different ways of forming their value expectations, it becomes impossible to make any predictions concerning social macro-conditions and individual behavior.<sup>13</sup>

The most interesting deprivation theories *do* specify assumptions about how people form value expectations. Several famous theories of revolution assume that people form their expectations with reference to the conditions of the past. One example is Davies’ (1962:6) “J-curve” thesis: “revolutions are most likely to occur when a prolonged period of objective economic and social development is followed by a short period of sharp reversal”. People here form value expectations about the future with reference to the steadily improving conditions of the past. When economic decline sets in, the discrepancy between sinking value capabilities and rising value expectations creates relative deprivation. Other theories assume that people or groups form their value expectations primarily by comparing themselves to other people or groups in domestic society. Aristotle, for example, thought the main cause of revolution was the common people’s aspiration for more economic and political equality, and the elites’ aspiration for greater inequality (Gurr 1970:37).

---

<sup>12</sup> For examples of how to include psychological or norms-based benefits in rational choice models, see Fearon and Laitin (1999:23) and Muller and Opp (1986:474).

<sup>13</sup> For example, a person with his/her neighbors as reference point would not feel deprived about an economic crisis striking all, whereas a person having his/her own past condition as reference would.

Norms are at the centre of all deprivation theory. People do not feel deprived merely because they want something which they cannot get. If this were true, most people would be discontented most of the time, as they are not capable of getting all that they desire. People feel deprived when they believe that some actor has the responsibility for their not getting something which they think themselves *rightfully entitled to*. Scott (1976) emphasized this point in his analysis of peasant rebellions against the colonial state in Southeast Asia. He argued that the colonial state provoked several rebellions in this region, especially during the 1930s, because its policies conflicted with the “subsistence ethics” of the peasants. Taxation and rent extraction were the main infuriating issues. But Scott argues that it was not primarily the increasing tax burden *per se* which angered the peasants. Rather, it was the rigidity of the tax system, which imposed a fixed charge on every peasant without considering his ability to pay or his subsistence needs (Scott 1976:92-93).

I think the most plausible grievance explanations for the poverty-civil war relationship go along the lines of Scott (1976), and concern state-society relationships. A static condition of poverty is probably not sufficient to make people rebel. Neither does economic decline necessarily cause deep resentment against the government or the state. “Grievance rebellions” against the state should be most likely to arise when the state can directly be linked to the grievance. The first issue I suggest may provoke especially strong grievances in poor societies is *increasing taxation*. We know that in the period of state formation in Europe, “tax rebellions” were relatively common (cf., Mousnier 1970). According to Tilly (1973:221), “[t]axation was the prominent single issue in the large-scale rebellions during the European state-making of the sixteenth to nineteenth centuries.” Though there has been less focus on this issue in analyses of civil wars in poor post-colonial countries, we should consider the possibility of increased extraction playing a similar role there.

The propensity of taxes to cause grievances depends on the degree to which they conflict with the prevailing norms of society. Most likely, the conflict is often intense in societies with a “new state” engaged in an early phase of state building, where extraction is raised from a low initial level. In Southeast Asia, for example, the colonial state represented a radical break with the extraction policy which people were accustomed to under the pre-colonial state (Scott 1976:53). We may hypothesize that state rulers trying to increase extraction in new post-colonial states similarly produced grievances and resistance. Most of these states had low levels of extraction at the time of independence, but increased this level quite dramatically in

the following years (Cohen et al. 1981:906). This argument may provide a possible explanation for the per capita income-civil war relationship because most of the post-colonial countries were, and still are, poor.

I try to clarify the argument with a causal diagram in *figure 1*.<sup>14</sup> In the figure, I consider a simplified version of the argument, which holds that the poverty effect is *indirect*. This presumes that poverty *causes* lower levels of extraction. It is very difficult to assess the truth content of this presumption. There are reasons to believe that at least parts of the relationship between poverty and extraction level is due to a causal effect from poverty to extraction level. With increased per capita income, the taxable base goes up because a lower proportion of income is devoted to subsistence (Fauvelle-Aymar 1999:403). However, it is possible that the causal effect may partly go in the other direction, from levels of extraction (indicating state strength) to per capita income. The relationship may also be partly spurious; the fact that state-building and economic development has been parallel historical processes may contain an element of coincidence. Thus, the argument may imply that the poverty-civil war relationship is partly spurious, and partly indirect, going through extraction level.

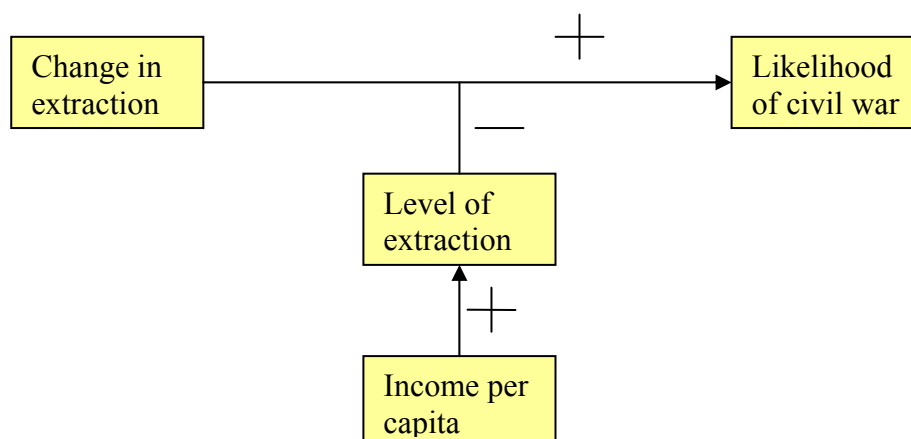
*Figure 1* shows that *change in extraction* is the primary causal variable in this argument.<sup>15</sup> Its positive causal effect on *risk of civil war* is larger the lower the *level of extraction*. This conditionality is captured by a line with a minus symbol going from *level of extraction* to the arrow showing the effect of *change in extraction*. As explained above, I have for simplicity considered the relationship between *per capita income* and *level of extraction* to be unidirectional and causal. The model considers the causal effect of *per capita income* to be indirect, passing through *level of extraction*. The indirect causal effect induces correlation between *per capita income* and *risk of civil war*. The model may thus explain why studies find a direct negative effect of *per capita income*: they have failed to control for *change in extraction* and *level of extraction*.

---

<sup>14</sup> The causal diagrams follow the template of Skog (2004:275-276). For simplicity, I have ignored the causal mechanisms – in this case grievance – in the figures.

<sup>15</sup> When discussing causal structure I write variable names in italics.

**Figure 1.** Causal diagram showing the grievance argument that increased extraction in states with initially low extraction levels explains the poverty-civil war relationship.



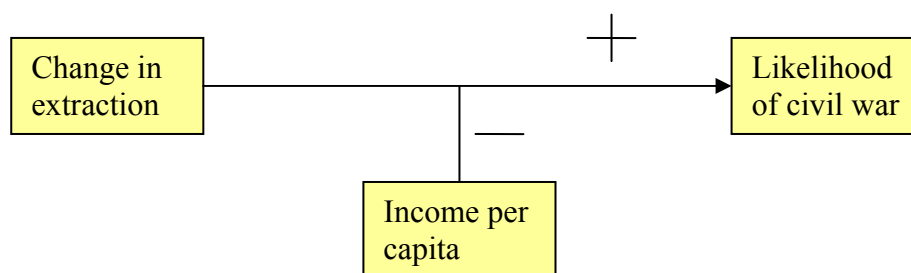
Explanation of symbols: Arrow = Causal effect  
Line meeting arrow = Interaction / Conditional effect

A less extreme version of the argument holds that only a part of the effect of *per capita income* on the *risk of civil war onset* goes indirectly through *level of extraction*, and the other part is direct and causal. I derive a testable hypothesis on the basis of this more plausible form of the argument. The logic of this derivation follows Skog (2004:276).

- H1:** (a) Increase of extraction from a low initial level of extraction is related to a higher risk of civil war onset.
- (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for the interaction between change of extraction and level of extraction.

I consider another way that increasing extraction can explain the poverty-civil war relationship. This builds on Scott's (1976) argument that tax increases sparked discontent in Southeast Asia mainly because they conflicted with people's "subsistence ethic"; they ignored people's ability to survive or maintain a livelihood. The argument can plausibly be generalized. Increased taxation should be more likely to threaten people's basic needs when it happens in very poor countries. As a general hypothesis, people are likely to consider this grossly unjust. *Figure 2* illustrates the causal logic of this argument in its simplest form. The causal effect of *per capita income* is due to it strengthening the causal effect of *change in extraction*. Thus, the poverty-civil war relationship can be explained by the interaction effect between *per capita income* and *change in extraction*.

**Figure 2.** Causal diagram showing the argument that increased extraction in poor countries explains the poverty-civil war relationship.



Explanation of symbols: Arrow = Causal effect  
Line meeting arrow = Interaction / Conditional effect

This argument is also more plausible if we allow for parts of the total effect of per capita income to be direct and negative. I derive the following hypothesis:

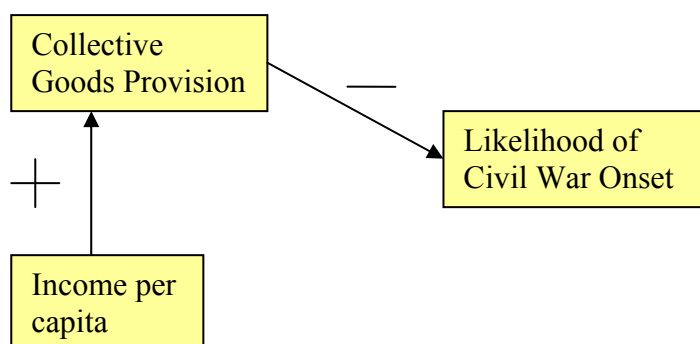
**H2:** Increase of extraction is related to a higher risk of civil war onset, and the effect is stronger the poorer is the country.

A second economic source of grievances against the state in poor countries may be lack of *public goods provision*. Contract scholars often emphasize that the level of public goods provision must correspond to the level of extraction to keep a viable social contract between state and citizens (Azam 2001). Discontent may thus arise if the state does not keep its part of the contract by providing a certain amount of public goods. Holsti (1996:109) argues that this may not apply in many new states: “In most post-1945 states (but not the post-Soviet republics), the idea of state responsibility to provide welfare beyond education, though prominent in the development literature, is not yet widespread.” Azam (2001:442) disagrees: “the occurrence of civil conflict in Africa is intimately related to the failure of governments to deliver the type of public expenditure that the people want, i.e., with a strong redistributive content such as in health and education.”

I will not argue that one of the two views is correct. Azam’s argument, which refers to what “people want”, is not entirely consistent with relative deprivation theory, however. Relative deprivation is not created because people do not get what they desire, but because they do not get what they believe they rightfully deserve. A proper relative deprivation argument needs to assume that people (also in Africa) have norms saying that the state is responsible for providing at least a modicum of public goods, such as education and health. We must keep in

mind that the argument rests on a very uncertain assumption. *Figure 3* lays out the argument in a simple form, where the entire causal effect of *per capita income* on the *risk of civil war* is *indirect*, going through *public goods provision*. It also rests on the unrealistic assumption that there is no reciprocal causal effect going from *public goods provision* to *per capita income*.

**Figure 3.** Causal diagram showing the argument that the causal effect of income per capita on risk of civil war goes through public goods provision.



Explanation of symbols: Arrow = Causal effect

I allow for the possibility that a minor part of the causal effect of *per capita income* is direct when stating the hypothesis:

- H3:** (a) A lower level of public goods provision, such as public education, is related to a higher risk of civil war onset.
- (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for public goods provision.

## 3.2 The state weakness explanation

*“State formation and consolidation has everywhere extracted horrendous human costs. Despite this, in the modern world it seems that only not having a state is worse than having one”*

Mehta (2003:106)

### 3.2.1 The structural state capacity argument

Fearon and Laitin (2003a) argue that the main factors determining the risk of civil war are the structural conditions that favour *insurgency*, defined as “a technology of military conflict characterized by small, lightly armed bands practicing guerrilla warfare from rural base areas”



(Fearon and Laitin 2003a:75). They hold that the motives and goals of rebels fighting civil wars are many and various. Still, rebels usually apply the same means to reach their goals: guerilla warfare. Thus, the key to understanding where and when civil war occurs, according to Fearon and Laitin, is not found by analyzing people's motivations to rebel.<sup>16</sup> They rather suggest a focus on *the conditions that make insurgency a feasible and attractive strategy for dissidents*. Most important of these conditions, according to the authors, are the government's police and military capabilities, and the reach of state institutions into rural areas. Correspondingly, they argue that a higher per capita income is associated with a lower risk of civil war primarily because wealthier countries have stronger states with larger administrative control over their peripheries (Fearon and Laitin 2003a:80).

The state capacity argument finds much support in the literature on insurgency and counter-insurgency. In his classic historical study of guerrilla warfare, Laqueur (1977:389-392) argues that one major failure of the scholarly literature on civil wars has been its relative neglect of analyzing the state and government side. He argues that guerrilla movements can only thrive in the face of a government that does not have *the capacity* or *the will* to use its full strength in the struggle. The lack of will and resolve is most likely in liberal regimes, he contends, where political concerns may prevent a government from using all available means to crush the rebels. More relevant for our concern is the importance he adds to state coercive capacity: "If government control and coercion is really efficient, a guerilla movement cannot possibly develop as the Communist and Fascist experience has shown" (Laqueur 1977:390).

Leites and Wolf (1970) also highlight state capacities in their influential book, *Rebellion and Authority*. In line with Fearon and Laitin (2003a), they dismiss what they call the "hearts-and-minds" approach to analyzing rebellion, which focuses on popular sympathies and economic conditions. Rather, they suggest a "systems view" of rebellion, focused on rebels' opportunities, costs and benefits, and the authorities' possibilities to affect these. Coercive capacity plays a prominent role in their theory. They do not emphasize regular military capabilities, however, but rather the state's ability to gather information about the rebels' location and supply routes. Accordingly, "[m]obility, reconnaissance, police (rather than military) intelligence, a capacity for operating in small units, and police and paramilitary forces are the important military elements for deterring or meeting the threat of R [rebellion] –

---

<sup>16</sup> Much of the literature on revolution and rebellion has focused on this question (Laqueur 1977:386-389).

not armor, artillery, jet aircraft, and large centralized operations by large divisional units” (Leites and Wolf 1970:154).

Kalyvas (2006) takes a similar analytical approach in his study of violence in civil war. Like Leites and Wolf (1970), he argues that coercive capacity and control is more important than the “hearts and minds” of the population. On his theory, most people will consider the costs and benefits, especially in terms of security, of supporting the rebel or government side in the war, and choose what best promotes their personal interest. In localities where the rebels can more effectively than the government sanction people supporting the other side, most people will support the rebels. Hence, he argues that “military resources generally trump the population’s prewar political and social preferences in spawning control” (Kalyvas 2006:111). His micro-level studies of the Greek civil war provide support for this contention (Kalyvas 2006:246-328).

Kalyvas’ theory of popular support is relevant for the study of civil war onset. Most students of guerilla war emphasize that collaboration of the local population is crucial for the rebels’ survival. T.E. Lawrence suggested that “Rebellion (...) must have (...) a population (...) sympathetic to the point of not betraying rebel movements to the enemy. Rebellion can be made by two percent active in a striking force, and 98 percent passively sympathetic” (cited in Leites and Wolf 1970:11-12). If control determines collaboration, rebellions should hardly be viable in countries where the state has a strong administrative and military presence throughout its territory.

Herbst (2000a; 2000b; 2004) has in various works on the state in Africa emphasized the importance of state coercive capacity and control for avoiding civil war. He suggests that “(...) the current rebel movements came about in part because the states in some African countries are so weak. Therefore, important leverage could be gained in ending these conflicts by increasing the coercive ability of the states that fight rebels” (Herbst 2000a:287). He particularly challenges many of the recent conclusions from economic studies of civil war, which recommend promoting development, diversifying the economy, stopping export of conflict diamonds, and often, cutting military expenditure to prevent civil wars (e.g., Collier and Hoeffler 2007; 2004). These studies neglect the essentially military nature of civil war, according to Herbst. He rather suggests that to prevent rebellions in Africa, we should first

promote the institutions responsible for order – primarily the military and the police (Herbst 2004:367).

### **3.2.2 The choices of weak-state rulers**

The structural state capacity argument has many scholarly adherents, and is intuitively strong. There are at least two objections to it, however. The first is its ignorance of the central actors' *choices*. The state is not only a structure which determines the opportunity for insurgency. It is also composed of actors making choices that crucially affect whether civil war will occur. As Gates (2004:5) puts it: "Civil wars originate, persist, and terminate with human decisions." Any argument which pretends to constitute a theory of civil war should therefore explain how macro-structures, like state capacities, relate to micro-level choices of the most important actors. A second objection may be that the theory is too static, and does not consider change. I discuss the first objection before returning to the second.

The applied study of Fearon and Laitin (2003a) does not lay out a comprehensive theory of the links between state strength and individual behaviour. In a related paper, Fearon and Laitin (1999) set up a game-theoretic model which explores the micro-foundations of insurgency. In this model, rational and self-interested members of a minority group decide whether to rebel or pursue their best alternative job. People consider the risk of rebellion, which increases with government counterinsurgent spending, as well as possible benefits and costs, which are many and various. The government simultaneously chooses how much to spend on counterinsurgency, with their goal being to minimize both the cost of rebellion and counterinsurgent spending (Fearon and Laitin 1999:22-23).

The impact of the structural context on these choices is analyzed through two parameters. The first is the conflict technology, or the *efficiency of government counterinsurgent spending*. The less efficient are the state's forces in fighting the rebels; the less will the government spend on counterinsurgency, *ceteris paribus*. Consequently, more people will choose to rebel because there is less risk and more to gain from it. There will thus be more rebellion and civil violence where the conflict technology favours insurgency (Fearon and Laitin 1999:26-28). The second parameter is the *per rebel cost of rebellion to the government*. Everything else held constant, the higher is the per rebel cost; the more will the government spend on counterinsurgency (Fearon and Laitin 1999:31-32).

These results may be less important than the analysis of what structural factors affect the model's two essential parameters. The authors suggest that the efficiency of counterinsurgent spending decreases with several factors, for example with rough terrain, a weak and poor central government, and the availability of support from neighbouring groups or states (Fearon and Laitin 1999:44). In their later study, Fearon and Laitin (2003a) do not dismiss any of these factors, but suggest a more state-centred explanation emphasizing state police and military capacity, as well as administrative reach into rural areas. The second parameter, the per rebel cost of rebellion to the government, can also be affected by several variables. In one interpretation, the authors see this parameter as the government's marginal cost of counterinsurgent spending. They suggest that this plausibly relates to the government's revenue: with less revenue, the marginal cost of counterinsurgent spending is higher.<sup>17</sup> Thus, they predict that poor governments will spend less on counterinsurgency, although facing a larger insurgent threat than wealthier governments (Fearon and Laitin 1999:31-32).

A more direct interpretation of the costs of rebellious activity to the government is not considered by the authors. They argue that it is difficult to find variables related to such an interpretation (Fearon and Laitin 1999:31). I think this issue deserves discussion. It seems plausible that *the economic cost of rebellion to the government is usually lower in weaker than in stronger states*. One reason is that weak-state governments of poor countries typically do not rely much on income taxes (Fauvelle-Aymar 1999:392). Thus, much revenue need not be lost when a rebel group takes control over a location. Second, weak states usually have little control over their periphery before an insurgency. Often, there is not much infrastructure which can be destroyed, and probably few local officials and institutions are in danger.

This argument receives support from Reno's (1998) account of "warlord politics" in African weak states. He describes how several rulers of African weak states, when their revenues plummeted after the end of the Cold War, turned to "warlord politics". They essentially gave up the pursuit of building a stronger state capable of serving the public good and providing security for all. Rather, they privatized several of the state's functions, scaled down the bureaucracy, and relied primarily on control of commerce through informal channels (Reno

---

<sup>17</sup> For the formal proof, see Fearon and Laitin (1999:31)

1998:1-3).<sup>18</sup> These weak-state rulers found that they could actually profit from disorder and collapsing authority through taking control over commercial hotspots with the help of foreign security and mining firms (Reno 1998:47). The rulers also perceived potential rebel groups much less of a security threat than the rivalling strongmen within the state itself.

Reno shows how weak-state rulers may under certain conditions find it in their personal interest to further weaken or totally dismantle formal state institutions. This may have been a rational survival strategy for the strained governments in the four countries he studies, but it certainly seems not to have been a good strategy for keeping the civil peace. In three of the four countries in his study, Sierra Leone, Liberia, and the Democratic Republic of Congo, but not Nigeria, civil war broke out in the 1990s (see war list in appendix). This fits well with the state strength explanation. As the state built down its administrative and coercive capacity, rebel groups seized the opportunities offered by the power vacuum. To sum up, state weakness may imply lower expected costs of rebellion and civil war for state rulers. This can mean that they choose not to take necessary but costly measures to avoid it.

In the above logic, state weakness “causes” civil war *indirectly*, as it induces weak-state rulers to make choices that increase the opportunities for rebellion. State weakness may also more directly increase the risk of civil war through its association with lower costs of war to governments. Fearon (1995) has aptly explained that war represents a “rationalist puzzle”. As long as both rival actors suffer costs of war, there exists a negotiated solution which makes both actors better off than with war. Rational actors should bargain on the basis of their relative strength and willingness to fight, and settle on the final distributive resolution that a war would bring, without suffering its costs. Obviously, the larger costs both actors expect to incur during war, the larger will the range of pareto-improving negotiated settlements be. State weakness may thus increase the risk of civil war by reducing this bargaining range.<sup>19</sup>

### **3.2.3 The choices of dissidents in weak states**

I have argued that state weakness should generally decrease the incentives for governments to make choices that secure the civil peace. What about dissidents or “potential rebels”? As

---

<sup>18</sup> This policy was most extreme in Zaire (now the Democratic Republic of Congo) under Mobutu, where the civil service shrank from 550,000 to 60,000 people, and government expenditure on social services reached zero in 1992 (Reno 1998:53, 154).

<sup>19</sup> Azam and Mesnard (2003), for example, emphasize the importance of costs of war for reaching a negotiated settlement (or social contract) between a dissident group and a government.

discussed above, rebels face larger security risks when facing a state with higher counterinsurgent capacity. The opportunities for rebels to profit during war should also be fewer when facing a stronger state. It becomes harder to seize settlements which the rebels can tax. Taking control over natural resources and trade routes is also more difficult the stronger is the state. But if the rebels seek not gains from the war economy, but rather to take control over the government's revenues, state weakness (or state poverty) could decrease the potential economic gain.

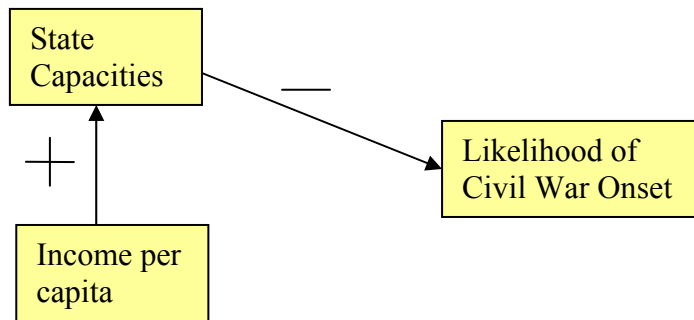
We do not know how many rebel groups are motivated by the prospect of capturing the state's revenues. However, if they have rational expectations, most rebel groups should hold a low estimated probability of taking over government power. This follows from the historical record of how many rebellions actually win an absolute victory. Laqueur (1977:405) found that no separatist guerilla movement had ever scored a decisive victory, and "[g]uerrilla war against domestic rulers has succeeded in the past – with one exception – only during a general war or immediately following it, with the collapse of central state power." Thus, only under the rare circumstances of a regular or "general" war, had rebel groups using guerrilla tactics taken hold of government at the end of the 1970s – Castro's revolution in Cuba being the only exception. By far the majority of rebellions do not capture state power. Thus, a larger state treasury should not significantly increase the expected benefits of most rebels. I therefore keep to the argument that state weakness generally *increases* the incentives to rebel.

### **3.2.4 Causal structure and a general hypothesis**

I conclude that the state capacity argument is coherent and convincing. It has usually been presented in structuralist terms, but it is also supported by considering the choices and incentives of the central actors. The degree to which this argument can explain the poverty-civil war relationship is uncertain, however. We should start by suggesting a clear causal structure. This is not easy because the studies reviewed do not explicitly account for how state capacity relate to poverty. Fearon and Laitin (2003a:80) argue that higher per capita income is a *proxy* for state capacities, which merely implies that they are highly correlated. The authors do not specify whether this is because poverty causes state weakness, visa versa, or whether there is also a coincidental, non-causal element involved. Thus, we cannot tell whether the per capita income effect on risk of civil war onset is hypothesized to be spurious or indirect, going through state capacities. I will not discuss this at length here, mainly because it does not matter for the hypotheses and the empirical tests.

The causal diagram in *figure 4* shows a simplified version where *per capita income* causes increased *state capacities*, and not visa versa. The entire causal effect goes from *per capita income* through the intermediate variable, *state capacities*. This induces the correlation between *per capita income* and risk of civil war, which is not due to any direct causal effect.

**Figure 4.** Causal diagram showing the argument that state capacities explain the poverty-civil war relationship.



Explanation of symbols: Arrow = Causal effect

I take a less extreme causal model as a basis for deriving the hypothesis, and allow for *per capita income* to have some direct causal effect.

- H4:** (a) Higher state capacities are associated with a lower risk of civil war onset.  
 (b) When controlling for state capacities, the major part of the per capita income effect on risk of civil war onset disappears.

My discussion has shown that there is disagreement concerning which specific state capacities are most relevant for preventing civil war. Even less agreement concerns how to measure these.<sup>20</sup> I discuss this here, and arrive at some lower level, less abstract hypotheses for testing the state weakness explanation.

### 3.2.5 Fearon and Laitin's state capacity proxies

Fearon and Laitin (2003a) use per capita income as their primary measure for a state's overall capabilities. I have already discussed that this is a very rough measure which cannot help us

<sup>20</sup> The literature on this is large. See Jackman (1993), Migdal (1981:279-287), Organski and Kugler (1980), and Arbetman and Kugler (1997).

separate between the various explanations. Four additional variables are considered as proxies for some form of state capacity in their study. First, “*new state*”, denoting whether the state newly reached independence, seems to be interpreted as a proxy for counterinsurgent capacity. Newly independent states are more prone to civil war because they have suddenly lost the military backing of the imperial power, and their own military is new and untested, according to Fearon and Laitin (2003a:81). This may be correct, but it is not difficult to come up with other relevant theoretical variables associated with new states. Sambanis (2004b:837) suggests that it may be causally related to ethnic fractionalization, and that ethnic fractionalization is a theoretically more interesting concept. It is also possible that newly independent countries are conflict-prone due to high political tensions. Thus, I find this measure problematic, and choose not use it for testing the state weakness explanation.

Second, whether fuel exports exceed one-third of export revenues (“*oil exporter*”) is thought to proxy a weak state apparatus. This is due to the infamous political “Dutch disease”; oil states have less need to extract taxes, and thus invest little in administrative institutions (Fearon and Laitin 2003a:81). I find also this proxy problematic because it is plausibly related to other explanations. Particularly, it may reflect grievances caused by the poor governance of many rentier states (Humphreys 2005). I therefore do not use this variable to test the state capacity explanation.

The final two variables that Fearon and Laitin use as proxies for state capacity are “*political instability*” and “*anocracy*”. *Political instability* measures whether the country had a large change in the type of political regime in the preceding year. Fearon and Laitin (2003a:81) interpret this as indicating “disorganization and weakness and thus an opportunity for a separatist or center-seeking rebellion”. In my view, change of political regime is not a good measure for the general strength or organizational capacity of the state. It might clearly reflect political instability, and possibly also an armed conflict in the making, but this is not equal to state capacity as defined here. The other variable, *anocracy*, is even more distant to my concept of state capacity. This variable marks regimes that can neither be characterized as full-blown democracies or autocracies. It is far from certain that *anocracy* indicates high political contestation, and that this in turn indicates state incapacity, as Fearon and Laitin (2003a:81) argue.



We thus need better indicators to test the state weakness explanation. The literature on state strength and insurgency suggests that all these capacities are relevant: economic strength and capacity to regulate economic relationships, counterinsurgent capacity, administrative capacity, ability to broadcast power and penetrate rural society, and “political capacity”, understood as the extent to which the state manages to achieve voluntary compliance from its citizens. Unfortunately, despite much effort, we lack good measures for all these capacities. As Migdal (1988:279) put it, “ascertaining the differences among states in their capabilities or the degree of social control they have exercised in their societies has been about as difficult for social scientists as actually establishing social control has been for state leaders.”

The lack of data prevents me from moving forward deductively from each type of state capacity to its measurement. Given the data available, the most transparent approach is to take the indicators as my starting point, and from there discuss their relation to the theoretical concept. I suggest four indicators: *military personnel per capita*, *road density*, *government share of consumption*, and the *tax ratio*.<sup>21</sup>

### **3.2.6 Military personnel: counterinsurgent capacity**

The capacity of the state to fight insurgents is central to the state weakness explanation. As shown in the above discussion, there is disagreement concerning which capacities are most important for counterinsurgency. Most researchers agree that police forces, paramilitias, intelligence, and the training and equipment of the police and military is important (cf., Leites and Wolf 1970:154; Herbst 2004). Unfortunately, we do not have sufficient cross-national data for these variables. I concur with Herbst (2004) that the best available measure seems to be *military personnel*. The importance of the military for fighting insurgents is disputed. I would think, however, that when an insurgency approaches the level of civil war, the capacity of all armed forces would matter.<sup>22</sup> Despite its shortcomings, it should be a more direct and better proxy than per capita income, which Fearon and Laitin (2003a) use.

**H4.1: (a)** The number of military personnel per capita is negatively related to the risk of civil war onset.

---

<sup>21</sup> Notably, the indicators need not be highly correlated as they may refer to different types of state capacity. I am not considering a reflexive measurement model, where state capacity *causes* the indicators.

<sup>22</sup> Also, as many scholars on African politics point out, in many of the countries that have experienced civil war, there are very few external threats (Herbst 1990). We therefore have reason to believe that a large proportion of the military forces are used to handle internal threats.

(b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for military personnel per capita.

### **3.2.7 Road density: state reach into rural society**

Herbst (2000b) has emphasized the importance of roads and infrastructure for the broadcasting of state power in Africa. He argues that the low density of roads, which effectively cuts large areas off from the capital, is at the heart of many African states' weakness. Fearon and Laitin (2003a:80) explicitly state that they see per capita income as a proxy for the degree to which terrain is "disciplined" by roads and rural society "penetrated by central administration". Their argument holds that road stocks is also a relevant *ecology indicator*. It does not only reflect the degree to which central administration is broadcasted, but also the degree to which the physical environment is conducive to guerilla warfare. Thus, this indicator cannot be used to separate between the state capacity and ecology explanations.

**H4.2:** (a) A lower road density is associated with a higher risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for road density.

### **3.2.8 Government share of consumption: economic resources**

A shortcoming of several proxies of state strength, such as per capita income, is that they capture the resources available in society, and not necessarily the state's ability to extract and control these. The latter can better be captured by measures of the relative size of the public sector to the private sector in the economy, and the ability of the state to collect taxes. Several researchers have suggested that state strength is closely associated with the relative size of the public sector (e.g. Carnoy 1984:3). It seems intuitively correct that the larger the proportion of economic resources in society controlled by the state administration; the greater the power of the state. These economic resources can be used to strengthen the bureaucracy, build roads and infrastructure, strengthen the police and military, or buy off political opponents or would-be rebels; all purposes that should decrease the risk of rebellion.

The *government share of consumption* is a relevant measure because it reflects the economic muscles of the state compared to other domestic social forces. It should not be interpreted as a proxy for the state's capacity to regulate and control economic relationships, however. States can effectively do this without having a large public sector. Japan, for example, has a small

government share of consumption, but the Japanese state has shown a remarkable ability to intervene in the economy (Jackman 1993:57-59). Conversely, many oil-rich countries illustrate that the state can be relatively incompetent even though it has a large public sector. Thus, the variable is probably not a good proxy for overall state capacity. I interpret it simply as an indicator of the state's available resources, which in several ways can be used to prevent rebellion.

**H4.3: (a)** A higher government share of consumption is associated with a lower risk of civil war onset. **(b)** The per capita income effect on risk of civil war onset is considerably weakened by controlling for government share of consumption.

### **3.2.9 Tax ratio: administrative reach and capacity**

Following Organski and Kugler (1980), several researchers have suggested that government revenues and extraction can say much more about a state's capacity than what I suggest above.<sup>23</sup> Arbetman and Kugler (1997:2) propose the grand vision that taxation data can be used to construct a measure which captures "political capacity" of the state just like GDP captures economic output. I argue that this vision is still very far from reality. Existing measures are not "exact indicators of governmental presence", as Organski and Kugler (1980:74) argue. I think extraction measures are interesting indicators of state strength, but call for more sober interpretations than what several researchers in this tradition offer.

The *tax ratio*, total tax revenue as a proportion of GDP, may capture other aspects of state strength than *government share of consumption*. Acquiring revenues through taxation is a challenging task for any state. High levels of tax collection require a highly capable state in several respects. Levi (1988) holds that successful extraction necessitates the "quasi-voluntary compliance" of taxpayers. That is, people must to some extent voluntarily choose to pay because they consider the taxes legitimate, but they should ultimately be deterred from noncompliance. Thus, to extract much revenue, the state must be ideologically and politically strong, as well as capable of detecting and punishing defectors. The latter requires both large administrative and coercive reach and capacity. According to Herbst (2000b:113), "[t]here is no better measure of a state's reach than its ability to collect taxes. If a state does not

---

<sup>23</sup> They can perhaps be called the modern followers of the "fiscal sociology" tradition created by Schumpeter and Goldscheid (Moore 2004).

effectively control a territory, it certainly will not be able to collect taxes in a sustained and efficient manner.”

How well the tax ratio measures cross-national variations in state strength is disputed. It is often argued that other factors than the administrative, political and coercive strength of the state affect the tax ratio (cf., Arbetman and Kugler 1997). Economic and technological differences probably create large variation in *potential tax ratio* between countries. In very poor countries it is probably not possible even for an administratively and politically strong state to extract a high proportion of total output. This is a weakness of the measure. Still, I find it very reasonable to believe that a higher tax ratio is associated with a larger administrative apparatus and more state presence in society. If this is partly due to level of economic development, it is not a large problem for my purpose. That would fit well with Fearon and Laitin’s argument that per capita income proxies state capacity.<sup>24</sup>

**H4.4: (a)** A higher tax ratio is associated with a lower risk of civil war onset.

**(b)** The per capita income effect on risk of civil war onset is considerably weakened by controlling for tax ratio.

I go on to discuss two other issues related to the state weakness explanation: *weakening states* (often named “state failure”) and *state building*. These may also provide explanations for the poverty-civil war relationship.

### **3.2.10 Weakening States**

Most of the historical and sociological research on the state and collective violence has focused on change in a system over time. Revolutions have usually been explained by the domestic societal changes preceding them (Coleman 1990:468). This approach often failed to consider differences in stable structural conditions which make some countries more prone to civil conflict than others. The recent quantitative studies of civil war may, on the other hand, may have tended to ignore change. They generally focus more on variations in structural variables across countries than over time. Regarding state capacity, much literature suggests that change should be carefully considered. After the Cold War much has been written about

---

<sup>24</sup> See chapter ??? for a more extensive discussion of alternative measures.

the instability associated with a decrease in the state capacity of weak states, leading to “state failure” or even “state collapse” (e.g., Rotberg 2004).<sup>25</sup>

Before the end of the Cold War, not much was written about states declining in strength or resources. The general trend until the late 1980s was that of states increasing their power. Carnoy (1984:3) held that the state “(...) has grown increasingly important in every society, from advanced industrial to Third World primary-goods exporter, and in every aspect of society (...)”. After the Cold War, however, several weak-state governments lost the aid that they had received from their superpower patrons. For example, US aid had formed a large proportion of the revenues of Mobutu’s government in Zaire and Doe’s government in Liberia. The sudden stop of aid flows sent these governments into a revenue crisis (Reno 1998:11). Both countries experienced civil wars during the 1990s.

There are good reasons to hypothesize that a sudden fall in the revenues of weak states should increase the risk of civil war. First, it should increase the risk through the state weakness-mechanism discussed above, giving potential rebels both larger opportunities and economic incentives to rebel. Second, it may increase the risk of rebellion through relative deprivation. Citizens and state employees may judge government performance on the basis of past performance. When revenues drop, the government may not be able to perform according to people’s expectations. This may create relative deprivation, which increases motivations to rebel. Thus, the hypothesis below cannot distinguish between the grievance and state capacity explanations. Third, as Reno (1998) argues, weak-state rulers may be more likely to choose warlord politics when facing a revenue crisis. This can in turn increase the risk of rebellion and civil war.

This may appear to be the exactly opposite prediction of the first grievance hypothesis, *H1*, which holds that increased extraction in states with low initial extraction level increases the risk of civil war. It is not necessarily so. Total state revenues consist of more than tax revenues. Importantly, it includes grants and aid. A poor, dependent state that suddenly loses much aid may experience a large decrease in total revenues while the extraction level is constant.<sup>26</sup>

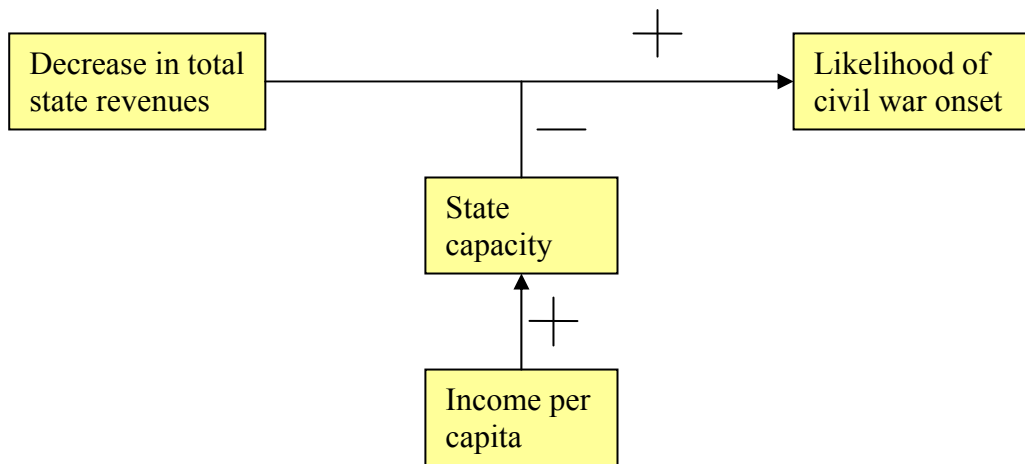
---

<sup>25</sup> See also the special issue of *Development and Change* (vol.33, nr.5, 2002) on “State failure, Collapse and Reconstruction”.

<sup>26</sup> I use *change in government share of consumption* as an indicator of revenue change (see paragraph 5.4.3). This measure shows a surprisingly low correlation with *change in tax ratio* ( $r = 0.05$ ) in my dataset.

The causal diagram in *figure 5* illustrates a simple version of this argument. Lower state capacity gives a larger positive effect of decrease in revenues on the risk of civil war. The finding of an income per capita effect is in turn explained by its positive relationship with state capacity. Thus the per capita income effect is thought to be both *indirect*, going through state capacity, and *conditional* on decreasing revenues.

**Figure 5.** Causal diagram showing the argument that decrease of revenues in countries with weak states explains the poverty-civil war relationship.



Explanation of symbols: Arrow =Causal effect  
Line meeting arrow = Interaction effect / Conditionality

I adopt a more plausible version of the argument when deducing a hypothesis, and allow for a considerable part of the GDP per capita effect to be direct and causal:

**H5: (a)** Decreasing government revenues in weak states are related to a higher risk of civil war onset.

**(b)** The per capita income effect on risk of civil war onset is considerably weakened by controlling for an interaction term between state capacity and change in revenues.

### 3.2.11 State building

The relationship between state building and risk of civil war is relevant to my first research question. Poor countries may be more prone to civil war not because of their low income levels or state capacity, but because they are going through a turbulent phase of state building. Holsti (1996:79) has argued that “[w]ar in the latter half of the twentieth century has been

primarily a phenomenon to (1) create states in the Western image, and (2) to try to hold them together after their creation.” Similarly, Cohen et al. (1981:901) argue that “violence in these [new] states is an integral part of the process of accumulation of power by the national state.” I have discussed the issue of state building and increased extraction from a grievance perspective. This section on the other hand looks at conflicting *interests* between the state and other domestic social forces.

Tilly (1985:182) concluded that state builders in Western Europe carried out four different activities: First, they *waged war* against external rivals whose power base was outside the territory they sought to control. Second, they engaged in *state making*, which meant subduing or eliminating internal rivals to create a monopoly of domestic violence. Third, they *provided protection* to those parties that supported their rule. Finally, in order to finance the first three activities, they *extracted resources* from the population within their territory.

Tilly pointed out that not only the first, but also the second task - achieving the internal monopoly of violence - was a violent affair:

“(…) the building of states in Western Europe cost tremendously in death, suffering, loss of rights, and unwilling surrender of land, goods, or labor (....) The fundamental reason for the high cost of European state building was its beginning in the midst of a decentralized, largely peasant social structure. Building differentiated, autonomous, centralized organizations with effective control of territories entailed eliminating or subordinating thousands of semi-autonomous authorities (....) Most of the European population resisted each phase of the creation of strong states” (Tilly 1975:71).

Some authors, including Tilly (1992), warn against drawing too many inferences from the European state-making experience to that of the developing world after colonialism. Bates (2001:82-83) emphasizes that the different international environment facing post-colonial state rulers has made state formation there different from the Western European. First, the new post-colonial states did not face large external threats. Thus, state builders of post-colonial countries have generally not engaged much in Tilly’s first activity – “war making”. Second, the sources of foreign capital were much more abundant for the post-colonial state builders than for the early European ones. Therefore, post-colonial state builders have also engaged much less in the fourth activity – taxation of its population – than the early European

state builders. Bates (2001:83) argues that this has created fewer incentives to forge liberal political institutions in the developing world.

Despite these differences, the rulers of post-colonial states have faced some challenges similar to those of early European state-builders. Most important perhaps is the challenge posed by rival domestic strongmen and social organizations. Post-colonial state leaders *have* needed to engage in Tilly's second activity – trying to subdue internal rivals. Cohen et al. (1981:902) suggest that this has often been a violent process: "Increasing central state claims for resources—for the material means of state-making and domination—intrude into and compete with preexisting structures of rights and obligations which tie those resources to subnational collectivities and/or "polities". Conflict, resistance, and violence are, as they were in Europe, often the result." Migdal (1988) also emphasizes that various chiefs, landlords, and their social organizations have many places put up strong resistance against an offensive state: "Strong Third World societies, then, are not mere putty to be molded by states with sufficient technical resources, managerial abilities, and committed personnel" (Migdal 1988:35).

Cohen et al. (1981) think we must separate between different phases of state formation. The first phase, involving "primitive central state power accumulation", is the violent one. During this phase in Europe, state leaders had to fight rival princes and lords to achieve something resembling a monopoly of violence, and often faced peasant rebellions over increased taxation. After the first phase, however, state building rather serves to *prevent* civil violence, according to Cohen et al. (1981:901-905). They argue that the state has then become strong enough to deter or crush s armed challenges. A challenge posed by this theory is to decide when and where a phase of primitive state power accumulation is taking place. Cohen et al. (1981:901-902) do not give a precise answer, but suggest that this phase lasted well into the 20<sup>th</sup> century in Europe. This is disputed. In Rokkan's (1975:570) model of European state formation, for example, the first phase of state-building - "penetration" - lasted only from the High Middle Ages to the French revolution.

Determining in retrospect when the old European countries went through the phase of primitive state power accumulation is difficult enough. It is probably more difficult to determine whether and when weak states of the post-World War II period went through this phase. Ayoob (1996:38) concurs with Cohen et al.'s theory, and holds that most developing countries today are at the same stage of state formation as Western European countries were



in the 16<sup>th</sup> to 18<sup>th</sup> centuries. He argues that as the old European states, the new states of the developing world need to use violence at this stage. They need “lots of time and a relatively free hand to persuade, cajole, and coerce the disparate populations under their nominal rule (...)” (Ayoob 1996:40).

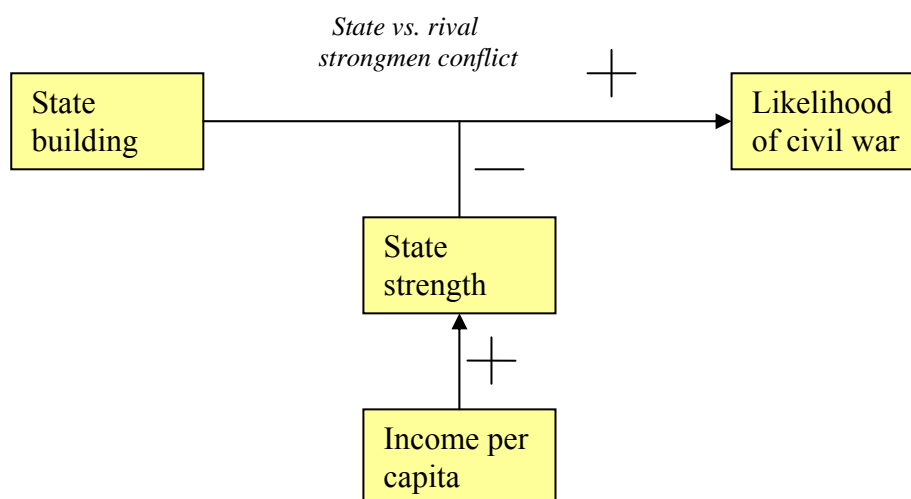
Ayoob also highlights the difference between the present era’s state formation in developing countries and the European process. First, today’s state builders in the developing world are trying to build their states in decades, whereas state making in Europe unfolded over at least three to four hundred years. This also means that the different phases of state formation, and basic tasks of the state makers, come simultaneously. Second, the modern state builders, unlike the old European ones, are pressured by international norms and donors to perform the task of state making in a humane way. Third, most state makers today must handle the challenges posed by mass political participation (Ayoob 1996:41). The last two points suggest that new constraints have made a violent state building strategy more costly, and hence less likely.

I have not arrived at a clear theoretical conclusion regarding the relationship between state-building and risk of civil war in post-World War II weak states. Cohen et al’s (1981) hypothesis should be tested, however. They suggest that “the lower the initial level of state power, the stronger the relationship between the rate of state expansion and collective violence“(Cohen et al. 1981:905). Cohen et al. (1981) use tax ratio as measure of state power. Their hypothesis is thus similar to the first grievance hypothesis, *HI*, which suggest that increased extraction leads to higher risk of civil war in countries with initially low levels of extraction. Cohen et al. (1981) and Ayoob (1996) point to a different causal mechanism than grievance, however. They hold that the causal link mainly goes through *conflicts of interest* between state-builders and rival strongmen or organizations in weak states. This provides another explanation for *HI*, which we cannot empirically separate from the grievance explanation.

*Figure 6* displays the argument in a causal diagram. In the grievance model (*figure 1*), “change of extraction” and “level of extraction” were the interacting causal variables. In this model, the theoretical concepts are substituted with “state building” and “state strength. The causal mechanism is not grievance, but the increased conflict of interest between the state and

rival strongmen. Like *figure 1*, this model considers the causal effect of per capita income to be indirect and conditional.

**Figure 6.** Causal diagram showing the argument that state building in weak states explains the poverty-civil war relationship.



Explanation of symbols: Arrow = Causal effect  
Line meeting arrow = Interaction effect / Conditionality

### 3.3 Ecological explanations

*“The essence of guerilla warfare lies in the fact that the guerilla can hide in the countryside and this, quite self-evidently, is impossible to do in a city”.*

Laqueur (1977:ix)

The “ecological approach” in social science emphasizes that actors’ behavior is conditioned by the environment or milieu in which it is situated (Sprout and Sprout 1965). The environment investigated can be natural or man-made, and concerns not only geography and resources, but also demography and technology. In civil war research much attention has recently been given to geography and natural resources (cf., Buhaug and Gates 2002; Ross 2004). These factors are probably not the most relevant for the first research question - why poorer countries face a higher chance of civil war onset than wealthier countries. Demographical and technological factors are probably more interesting, as these may point to some significant general differences between less and more developed countries. I consider three ecological explanatory variables: Settlement patterns and stocks of roads and telecommunication infrastructure.

### **3.3.1 Pattern of settlement**

Kocher (2004:3-5) suggests that a country's pattern of settlement is the key variable determining its risk of civil war. Like Fearon and Laitin (2003a), he emphasizes that the vast majority of 20<sup>th</sup> century civil wars took the form of insurgencies. Kocher further holds that insurgency and civil war is mainly about state formation; the insurgents need to build a parallel "counter-state" capable of controlling populations and extracting resources to survive. The central argument is that insurgency, as a form of competitive state building, is only viable where irregular forces can achieve control over a significant population without directly or decisively fighting the state's military. An urban settlement structure does not allow this, according to Kocher (2004:4): "Towns and cities force conventional battle, render intensive population control and surveillance relatively easy (via curfews, searches, roadblocks, identity checks, etc.) represent concentrated assets the state has high incentives to protect, and facilitate the denunciation of clandestine rebels to state forces".

The difficulty of launching insurgency in urban areas seems theoretically well grounded. Students of insurgency and guerilla war emphasize that the insurgents usually are militarily weaker than the state, particularly at the beginning of insurgency (Herbst 2004:360). They therefore need to avoid direct and open battles with government forces. This is possible in environments where the insurgents can hide from government agents, and where the state cannot deploy its full military force (Laqueur 1977:392-393). Clearly, cities do not facilitate hiding for an army of several hundreds or even thousands of armed insurgents. Kocher (2004:27-28) also argues that the state can more easily apply its full coercive capacity in cities than in rural areas. This may be correct, but it does not seem as certain as the information and control argument. There are probably often significant limitations of a political or moral kind on the state's use of military force in city environments. Using massive force in densely populated areas could cause many civilian casualties, which will usually be costly in political, if not in moral terms.

While all of Kocher's (2004) arguments may not be equally strong, there are solid reasons to think that insurgencies do not thrive in urban areas. There is also a large scholarly consensus on this point (e.g., Fearon and Laitin 2003a:80; Gurr 1970:266; Kalyvas 2006:133). Kocher takes the argument further than most others, however, and challenges the state weakness explanation and economic explanations of the poverty-civil war relationship. He makes a controversial claim: "The rich and powerful states of the late 20th century *have* experienced fewer civil wars,

but not *because* they are rich and powerful. The most economically and politically developed states tend to have a very similar human ecology: they are highly urbanized. Conditional on a given level of urban-ness, however, rich and powerful states are no less likely, and may be somewhat more likely, to experience civil wars” (Kocher 2004:21).

The dismissal of the state weakness explanation seems to fit awkwardly with Kocher’s theory of insurgency and civil war. His theory emphasizes that insurgents must be able to hide from, but also challenge the state militarily through irregular combat. If the state had high administrative, police, and intelligence capacity, as well as large control over its entire territory and population, this should make insurgency difficult even in rural, village-based environments. The ecological and state weakness explanations are difficult to separate empirically because a high degree of urbanization is correlated with state strength (Kirby and Ward 1991:123-124). They may also be causally associated. Tilly (1992) stresses how cities played an important role in the formation of European states. They were “containers of capital” that state leaders made use of to eliminate internal and external rivals. Are states always stronger in more urban countries?

Herbst’s (2000b) account of how settlement patterns affect state building in Africa may illustrate that the relationship is more complex. He divides African countries into three categories according to the challenge posed by their political geography. Interestingly, he does not find that the environment of the “hinterland countries”, which have the largest areas of sparse or no settlement, the most challenging. These countries often have most of its population in a small area of high population density, which can be controlled by the state. The most difficult political geography displays several dispersed areas with high population densities. These population centers are not contiguous, and often large, scarcely populated areas separate them. This settlement pattern makes it very costly and difficult to extend control over the majority of the population. There is also the danger that the distant population centers can become strongholds of opposition to the central state. Herbst finds that the African countries with the most favorable political geography are quite densely populated. The highest population density is around the capital, with population densities decreasing as the distance from the capital increases (Herbst 2000b:145-154).

Herbst’s theory suggests that it may be possible to separate between Kocher’s ecological explanation and the state capacity explanation. State leaders may have troubles with taking advantage of the opportunities for extraction and control offered by densely populated areas if

these are spread over a large territory and the state's resources are strained to begin with. Distant urban areas may also be turned against the state by rival actors. As Laqueur (1977:393) has noted, guerilla warfare in the hinterland has its limitations. Guerillas need arms supplies, food, and recruits, which are best available in cities and towns. The most favorable setting for the guerillas might therefore be towns that have a low state presence and that are relatively inaccessible for the state's forces (Laqueur 1977:393).

Despite these objections, I agree with Kocher (2004) that the settlement pattern explanation is theoretically strong. I disagree with his contention that it renders all other explanations useless or irrelevant, however. In particular, the state capacity explanation should be complementary to his ecological account. The discussion has also pointed to the theoretical possibility that the degree of urbanization and state strength need not be perfectly correlated. The state *can* be relatively weak, although most people live in urban settlements. I derive the following hypothesis from the urban settlement argument:

**H6:** (a) A higher proportion of people living in urban areas is related to a lower risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for the proportion of people living in urban areas

### **3.3.2 Roads and telecommunication infrastructure**

Students of insurgency and guerilla war all emphasize that insurgents usually need to hide from state forces to survive. This is also the case for Fearon and Laitin (2003a:79-80): "If government forces knew who the rebels were and how to find them, they would be fairly easily destroyed or captured. This is true even in states whose military and police capacities are low." Guerilla strategists themselves highlight the need for hiding and avoiding places where the state has high access to information. Mao noted about the opportunities for the insurgents in China of the late 1930s: "When we are devoting ourselves to warfare in an open region, it is the (...) areas with a low cultural level, where communications are difficult and facilities for transmitting correspondence are inadequate, that are advantageous" (cited in Leites and Wolf 1970:133).

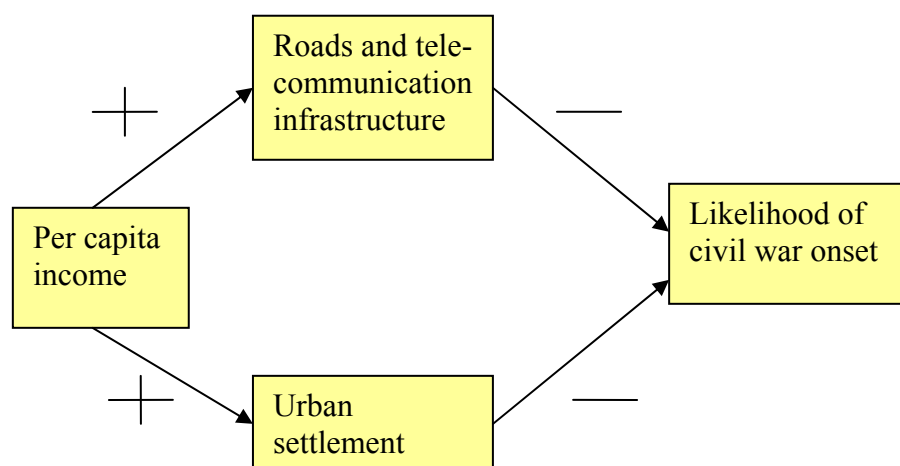
Leites and Wolf (1970) argue that access to information is of immense importance for the state to succeed in counterinsurgency. They think that developing countries are more prone to civil

war primarily because *the cost of information is higher* than in more developed countries (Leites and Wolf (1970:132). In poor countries, getting hold of information is difficult, it takes a long time, and the information you get is often unreliable. This ecological factor facilitates hiding for the rebels, and makes counterinsurgency difficult for the state.

Not only settlement patterns, but also the stock of infrastructure, such as roads and telecommunications, should affect the availability and cost of information. Fearon and Laitin (2003a:80) think that lack of road networks favors insurgency, but do not test this directly. Their hypothesis finds support in Kalyvas (2006:135), who states that proximity to central roads increases chance of people collaborating with the government in civil war. Laqueur's (1977) account of guerilla war may also point to the importance of roads and telecommunications. He finds that "the ideal guerrilla territory while relatively inaccessible should be located not too far from cities and villages" (Laqueur 1977:393). Thus, a good environment for insurgency should be one where a lack of roads and telecommunications makes some areas with relatively high population density disconnected from the capital and state control.

In *figure 7* I set up a simple causal structure where the entire per capita income effect on risk of civil war is indirect, going through the ecological variables.

**Figure 7.** Causal diagram showing the argument that ecological variables explain the poverty-civil war relationship.



Explanation of symbols: Arrow = causal effect

I formulate hypotheses on the basis of a less extreme version of the argument, allowing for some of the per capita income effect to be direct. The hypothesis concerning roads is already stated in the state weakness section (*H4.2*). The ecology and state weakness explanations both suggest that more extensive road networks decrease the risk of civil war onset.

**H7:** (a) Lower stocks of telecommunications infrastructure are related to a higher risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for telecommunication infrastructure.

### 3.4 The greed explanation

*"The true cause of much civil war is not the loud discourse of grievance but the silent force of greed"*  
Collier (2000a:101)

#### 3.4.1 Collier and Hoeffler's theoretical works

Collier and Hoeffler's (2004) "opportunity theory of rebellion" is based on economic models of conflict, pioneered by Herschel Grossman and Jack Hirshleifer. These models assume rational behaviour, and purely self-regarding preferences.<sup>27</sup> Moreover, rebellion is seen as a predatory and criminal activity. Rebels are thought to follow what Hirshleifer (2001:10-11) called the "Machiavelli theorem": no one will ever pass up an opportunity to gain by exploiting someone else. Collier and Hoeffler's (2004) central argument is that people have more to gain from rebellion when income opportunities in the peace economy are low. This translates into lower labour costs for rebel organizations, which gives better opportunities for rebellion. In their empirical analysis, per capita income, male secondary schooling, and the growth rate of the economy, are seen as proxies for the opportunity cost, or foregone income, of rebellion (Collier and Hoeffler 2004:569).

The economic opportunity theory is informally and briefly presented in Collier and Hoeffler (2004). The authors have in previous works tried to underpin their theory with formal models and more extensive discussions. In "On Economic Causes of Civil War" they set up an expected utility model of rebellion. A simple causal mechanism is suggested: "[w]ar occurs if

---

<sup>27</sup> Rational choice theory does not *necessitate* self-interested preferences. For more on the assumptions underlying various rational choice approaches, see Harsanyi (1986:83-89) and Bates (1983:134-135). For an overview of economic theories of conflict, see Garfinkel and Skaperdas (2007).

the incentive for rebellion is sufficiently large relative to the costs” (Collier and Hoeffler 1998:563). They hold that rebels’ objectives are either to capture the state or to secede from it, but do not consider secessionist rebellions much. The expected benefit of rebellion is thus the product of the probability of winning the war and the gains of victory, which depends on the state’s revenues.<sup>28</sup>

Per capita income affects three variables in Collier and Hoeffler’s (1998) early model. The higher per capita income; the larger is the taxable base of the economy. This means that the government’s counterinsurgent capacity increases with per capita income, translating into a smaller probability of rebel victory. On the other side of the equation, the gain in case of rebel victory also rises. The third variable affected is the opportunity cost, or foregone income, of rebellion. The lower per capita income; the lower is the opportunity cost of rebellion. In sum, the effect of per capita income on rebellion is theoretically ambiguous in this model because it enters on both the benefit and cost sides of the equation (Collier and Hoeffler 1998:564-565).

Collier (2000b) sets up a different economic model of rebellion, where rebels seek not to capture the state, but rather to loot natural resources on a continuing basis. This is akin to the “rebellion as crime” models of Grossman (1991) and Hirshleifer (2001). Collier (2000b:840) points out, however, that rebellion has different causes than other types of crime, particularly because it requires a much stronger military organization. He therefore treats rebellion as a “quasi-criminal phenomena”. The model is built around some central empirical findings, one being the strong inverse empirical relationship between per capita income and risk of civil war. Collier (2000b:848) finds this a “surprisingly difficult result to get into a simple model”.<sup>29</sup> In his new model, one source of GDP per capita’s ambiguity is removed, but another one comes in.

In this quasi-criminal model of rebellion a higher income level no longer increases rebels’ expected profits, because rebels do not seek the state’s revenues. However, the effect of GDP per capita on the expected utility of rebellion is still ambiguous, because the government is included as an actor in the analysis. The rebel organization and the government are seen as competing for labor.<sup>30</sup> Consequently, not only the rebel organization, but also the state should

---

<sup>28</sup> This differs from the later models of Collier and Hoeffler (2004) and Collier (2000b), where rebels do not seek governmental power, but rather to profit *during* civil war, primarily by appropriating natural resource rents.

<sup>29</sup> The ad hoc-theorizing of this study is surprisingly explicit.

<sup>30</sup> This is not an unrealistic assumption. In many civil wars, former rebel soldiers have turned over to the government army and vice-versa (Kalyvas 2006:98).



afford a larger army in low-income countries (Collier 2000b:844-845). It is therefore not certain that lower average incomes always makes it easier to form a rebel army more capable of fighting the state. The crucial question becomes: Why should the rebels draw more advantage of low labor costs than the government?

Collier (2000b) offers two arguments that may answer the question. First, he argues that the central task for explaining civil war is to account for the formation of a rebel army, not a government army. He reasonably asserts that the government will afford to keep an army under most conditions. The rebel organization may thus be more sensitive to labor costs than is the government. Lower average income is therefore a relative advantage for the rebels (Collier 2000b:841). Second, he points to the fact that government revenue increases more than proportionately in income (Collier 2000b:849). As the share of income devoted to subsistence decreases, the state can collect a larger proportion of national income (Fauvelle-Aymar 1999:403). Hence, the state gains advantage relative to potential insurgents from a higher income level. These arguments seem plausible. The first argument, however, rests on a relatively uncertain assumption. The second argument, on the other hand, diverges significantly from the economic motive explanation. Actually, it concerns *state capacity*, and serves to strengthen Fearon and Laitin's interpretation of the poverty-civil war relationship.

### **3.4.2 Other economic conflict models**

Collier and Hoeffler's explanation of the poverty-civil war relationship is not supported by the theoretical economic conflict literature. In most economic game-theoretic models, the level of income and the productivity of the economy have no impact on conflict behavior. Take Grossman's (1991) seminal conflict model. His game is between peasant families and a state ruler. The peasant families choose how much of their time they will spend on three income-generating activities: production, soldiering (for the state), and rebellion (against the state). The state ruler has three policy decisions: determining how much he will tax the peasants' income, deciding soldiers' wage level, and the number of soldiers. Both are assumed to be risk-neutral and choose the strategy which maximizes their expected economic gain. Fighting is assumed to be non-destructive, and the economy is closed, meaning that the ruler receives all his revenues from the peasants. In this game, income levels or the productivity of the economy does not affect the time allocated to conflict activity because the efficiency of labor time changes the return from production, soldiering and insurrection equiproportionately (Grossman 1991:912-917). Simply put, with increasing productivity, there will not only be more to gain from

producing, but also more state revenues to capture by rebelling and more to earn by serving in the state army.

Hirshleifer's (2001) classic model in "The Paradox of Power" similarly finds that productivity is irrelevant for the actors' decisions to produce or predate. He considers a closed economy with two actors, a richer and a poorer contestant, who can choose to fight over all the income produced in the economy. He finds that "an increase in overall economic productivity leaves the proportionate allocation of resources between producing and fighting unchanged" (Hirshleifer 2001:51). The central point is that increased productivity not only increases the return from production, but also the possible return from predation and conflict in these models, since income can be captured. Aggregate economic productivity and level of development neither affect the potential rebels' or the state rulers' utility calculus directly in other relevant game-theoretic models (e.g., Azam and Mesnard 2003; Skaperdas 1992; Azam 2006).<sup>31</sup>

### 3.4.3 Hypotheses from the greed theory

I have argued that the simple opportunity cost argument has a weak theoretical foundation. It has not been convincingly shown that lower incomes will generally favor the rebels vis-à-vis the state. *The effect of per capita income on the individual's decision to become a rebel, a government soldier, or pursue another job, is theoretically ambiguous.* This is a simple argument, but to my knowledge it has not been put forth in the quantitative literature on civil war. Collier's (2000b) ad-hoc arguments can not be *a priori* dismissed, however. Rebel recruitment may be more important than government soldier recruitment for explaining civil war, and rebel organizations may be more sensitive than the government to income levels for recruitment. I therefore want to test the hypothesis from the economic opportunity cost argument:

**H8:** The risk of civil war onset increases with the number of people in domestic society that have low economic opportunities in the regular labor market.<sup>32</sup>

---

<sup>31</sup> Acemoglu and Robinson (2001) is an exception. They suggest that lower aggregate productivity means a lower opportunity cost of revolution for the poor under non-democracy and of coup for the rich under democracy. They do not deduce this formally from their model, however. It is merely assumed at the beginning of the game (Acemoglu and Robinson 2001:941, 944).

<sup>32</sup> The hypothesis is generally stated. We derive more specific hypotheses below.

Collier and Hoeffler (2004:569) suggest three proxies for economic opportunity: GDP per capita, male secondary schooling, and the rate of economic growth. I agree that these are relevant indicators. GDP per capita reflects average incomes in the regular labor market. Thus, the greed explanation, unlike the other three explanations, holds that per capita income has a direct and negative causal effect on the risk of civil war onset. I may derive the following lower level hypothesis of *H8*.

***H8.1:*** Income per capita is negatively related to the risk of civil war onset even when controlling for the causal variables of the other explanations.

Particularly because the greed explanation is theoretically unconvincing, it is important to test as many of its empirical implications as possible. Income per capita does not uniquely capture economic opportunities for people in the labor market. Several other variables plausibly also affect economic opportunity. If the greed interpretation of the income per capita effect is correct, then other indicators of economic opportunity should also be related to a higher risk of civil war onset. Like Collier and Hoeffler, I test for *economic growth*, which is plausibly associated with new and better opportunities in the labor market. I also include *male secondary schooling* because it reflects educational opportunities for young males, the group from which most rebels are drawn.

***H8.2:*** Economic growth is negatively related to the risk of civil war onset.

***H8.3:*** Male secondary schooling is negatively related to the risk of civil war onset.

The problem with these three measures is that they do not uniquely indicate economic opportunities. They probably all capture state capacity to some degree. This is especially the case with per capita income, but it can also apply to male secondary schooling and economic growth. It has been argued that state capacity is related to economic growth (Englebert 2000). Also, richer states tend to spend more on education. As it cannot perfectly be controlled for state capacity, additional measures are needed to test the explanation. A very specific proxy for economic opportunity would be *unemployment rates*. Unfortunately, unemployment data have too low coverage across time and countries to allow for testing (World Bank 2006). I suggest two other indicators: *income inequality* and *level of extraction*.

Using *income inequality* as a proxy for economic opportunity is controversial. It starkly contradicts Collier and Hoeffler's (2004) interpretation of income inequality. They consider it a proxy for *grievances*, not for economic opportunity. I argue that this is misleading. A highly unequal society should clearly facilitate greed rebellions. The primary reason is that *higher income inequality at any given level of per capita income implies a larger pool of people offering cheap labor*. This argument follows a simple logic. I elaborate because it runs counter to Collier and Hoeffler's (2004) account.

Income inequality is usually represented by a Gini coefficient. A Gini of 0 means perfect equality and a Gini of 1 (with individual data) means that total income accrues to one person alone, leaving the rest with none.<sup>33</sup> Gini coefficients in the higher end of the scale thus imply lower income for a larger proportion of the population at a given level of GDP per capita. Consider an example: Brazil in 1995 had ca. 3 times larger GDP per capita than Ukraine in 1992 (World Bank 2006). Whereas income in Ukraine was quite equally distributed (Gini = 0,23), Brazil had a highly unequal income distribution (Gini = 0,60) (UNU/WIDER 2007). Consequently, 43.5 % of Brazil's population had an income below \$2 PPP a day, whereas less than 2 % of the Ukrainians lived for under \$2 PPP a day (World Bank 2000:230). *Ceteris paribus*, Brazil's economy should therefore be more conducive to "greed rebellion" than Hungary's, despite a much higher per capita income.

A second reason why inequality should facilitate greed rebellion is that unequal societies offer the opportunity of cheap rebel labor at the same time as there is high total income in the economy for rebels to capture. I find it strange that Collier and Hoeffler (2004) do not consider this. In other economic conflict models, like Hirshleifer's (2001) and Acemoglu and Robinson's (2001:949), larger inequality increases the economic motivation for rebellion. In Hirshleifer's model of a relatively anarchical power struggle between the rich and the poor, the poor will devote more effort to conflict the larger is the difference in incomes. He argues that in very unequal societies, the poor have a "comparative advantage in fighting". They should therefore specialize in conflict, and steal instead of produce (Hirshleifer 2001:54, 62). Similarly, in Acemoglu and Robinson's (2001:949) model of political transitions, the poor have more to gain by rebelling against a dictatorship of the rich if inequality is high. There are thus good reasons for deriving the following greed hypothesis:

---

<sup>33</sup> For a simple account of the logic of the Gini, see WIDER (2000). For calculations, see Damgaard (2003).

**H8.4:** Income inequality is positively related to the risk of civil war onset.

I finally suggest *level of extraction* as an indicator of economic opportunity cost. Tax burden should be negatively related to economic opportunities in the regular labor market. The larger proportion of people's income that the state collects in taxes, the lower is the foregone income of rebellion. Rebels, unlike regular workers, can escape taxation. Hence, a higher tax level means larger incentives to rebel, *ceteris paribus*. This in turn implies lower labor costs for rebel organizations.

This prediction has the advantage of going in the opposite direction of what the state capacity explanation predicts. Hence, this bears semblance of a critical test. High levels of extraction should be positively associated with state strength, as I argue below (Organski and Kugler 1980). State weakness theory therefore predicts a negative effect of increased extraction on the risk of civil war. However, Collier and Hoeffler think not only economic opportunities, but *also* higher state capacities reduce the risk of civil war. Therefore we cannot simply ascribe to their theory the prediction that the higher proportion of income that the state extracts, the higher the risk of civil war. The most satisfactory solution would perhaps be to estimate an "ideal level" of taxation, above which the effect of increased state capacities is offset by the effect of people's increased incentives to rebel instead of produce. This is clearly a difficult task, out of bounds for this thesis.<sup>34</sup>

A more feasible solution is to test for extraction level while controlling for state capacities that are strongly and specifically related to the opportunity for insurgency. Greed theory may imply that extraction level is positively related to risk of civil war when counterinsurgent capacity is held at a constant level. This is a disputable hypothesis, and it is also difficult to test due to measurement problems. With careful interpretation, though, I consider it worth testing.

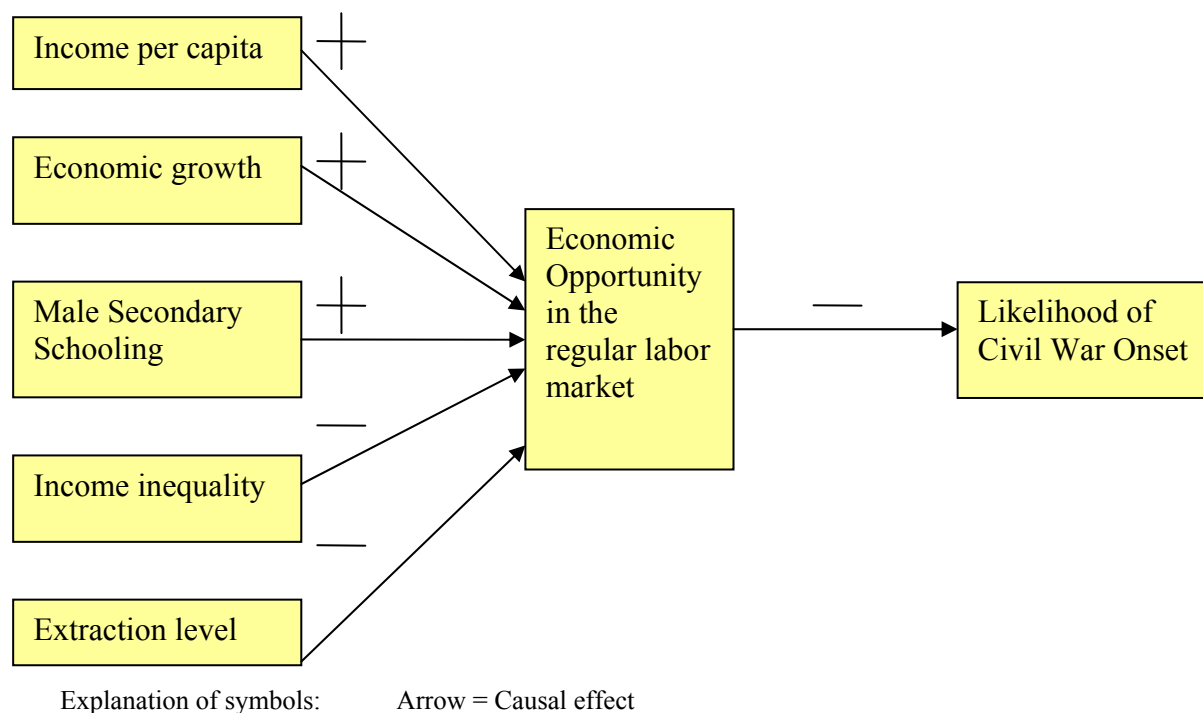
**H8.5:** Extraction level is positively related to the risk of civil war onset when controlling for state counterinsurgent capacity.

---

<sup>34</sup> See Edwards (1997) for a related discussion about the difficulties of estimating an "economic frontier", showing every state's possibility for extraction given economic and technological constraints.

Figure 5 shows the logic behind the greed hypotheses. The variables do not condition or explain the per capita income effect. They are different *formative* indicators of the same latent theoretical concept – economic opportunity in the regular labor market.

**Figure 8.** Causal diagram showing the greed / opportunity cost model.



### 3.5 Summary of hypotheses

I have suggested several hypotheses which can to a certain extent separate the types of explanations, and thus allow for testing. Clearly, not all of them correspond exclusively to one explanation. Still, they constitute an improvement from the prediction of an *unconditional* per capita income effect, which all the four theories may support. I believe my hypotheses are more strongly related to one or two of the explanations than the others. An overview of the hypotheses and the most closely corresponding explanations is given in *Table 1*.

**Table 1.** Overview of hypotheses and corresponding explanations.

Hypotheses	Explanations
<b>H1:</b> (a) Increase of extraction from a low initial level of extraction is related to a higher risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for the interaction between change of extraction and level of extraction.	Grievance + state weakness (states vs. rival strongmen account)
<b>H2:</b> Increase of extraction is related to a higher risk of civil war onset, and the effect is stronger the poorer is the country.	Grievance (less related to the states vs. strongmen account)
<b>H3:</b> (a) A lower levels of public goods provision, such as public education, is related to a higher risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for public goods provision.	Grievance
<b>H4.1:</b> (a) The number of military personnel per capita is negatively related to the risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for military personnel per capita.	State weakness
<b>H4.2:</b> (a) A lower road density is associated with a higher risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for road density.	State weakness + ecological explanation
<b>H4.3:</b> (a) A higher government share of consumption is associated with a lower risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for government share of consumption.	State weakness
<b>H4.4:</b> (a) A higher tax ratio is associated with a lower risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for tax ratio.	State weakness
<b>H5:</b> (a) Decreasing government revenues in weak states are related to a higher the risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for an interaction term between state capacity and change in revenues.	State weakness + grievance (possibly)
<b>H6:</b> (a) A higher proportion of people living in urban areas is related to a lower risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for the proportion of people living in urban areas.	Ecological explanation
<b>H7:</b> (a) Lower stocks of telecommunications infrastructure are related to a higher risk of civil war onset. (b) The per capita income effect on risk of civil war onset is considerably weakened by controlling for telecommunication infrastructure.	Ecological explanation

<b>H8.1:</b> Income per capita is negatively related to the risk of civil war onset, also when controlling for the causal variables of the other explanations.	Greed
<b>H8.2:</b> Economic growth is negatively related to the risk of civil war onset.	Greed
<b>H8.3:</b> Male secondary schooling is negatively related to the risk of civil war onset.	Greed
<b>H8.4:</b> Income inequality is positively related to the risk of civil war onset.	Greed
<b>H8.5:</b> Extraction level is positively related to the risk of civil war onset when controlling for state counterinsurgent capacity.	Greed



## 4 Statistical method

In this chapter I account for the method used to test the hypotheses. I begin with an argument for applying statistical analysis of large-N cross-national data, as well as an account of the logistic regression model. The central challenge for my analysis, multicollinearity, is then considered before I explain for the logic and structure of the hypothesis testing. Then I discuss methodological challenges and possible remedies for these. Finally I consider which control variables should be included in my analysis.

### 4.1 Logit analysis of cross-national time-series data

I investigate the hypotheses using logistic regression analysis of Fearon and Laitin's (2003a) civil war data. These data cover all independent countries with more than 500,000 inhabitants from 1945 to 1999.<sup>35</sup> This type of cross-national time-series analysis of civil war has boomed since the late 1990s (Gates 2002:1). Recently, James Fearon suggested that it may have depleted its usefulness largely due to the difficulty of testing the rival explanations of the per capita income finding (noted in Kocher 2004:50-51). I agree that there is a fundamental problem facing cross-national quantitative research which makes such testing difficult: the shortcomings of available empirical measures. Most of the measures with good data coverage across time and countries are not theoretically determinate – they are not highly valid measures of only one explanation. This is arguably the greatest problem facing all large-N research (King and Zeng 2001b:714).

I still believe there is more to be learned from analysis of cross-national time-series data. In this study I suggest several previously untested measures that I think can better separate the explanations. Some of these measures unfortunately cover a shorter period or fewer countries than per capita income, and some have significant reliability problems. I would argue, however, that they bring valuable information into the analysis. Testing them statistically can be meaningful provided that possible bias is carefully discussed. My study also contributes to this empirical literature by testing hypotheses concerning conditional, indirect, and spurious effects. To investigate the poverty-civil war relationship we should test how various explanatory variables are related both to the dependent variable and to per capita income. This has not been systematically carried out in this literature before. Analysis of large-N cross-

national data thus still holds a potential for increasing our knowledge of the puzzling poverty-civil war relationship.

The dependent variable (Y) in my analysis, civil war onset, is a dummy taking the values 1 (onset) and 0 (no onset). I want to estimate how the risk of onset is affected by changes in the independent variables. We can not assume that the relationship between the independent variables and the probability of onset is linear, however, as in standard OLS regression. Such a linear regression model would imply a possibility for predicting values of Y outside the 0-1 interval. To escape this problem, we may instead estimate the *log-odds* of onset for different values of the independent variables. The odds of onset for country  $i$  in year  $t$  is the probability of onset ( $P_{it}$ ) divided by the probability of no onset ( $1 - P_{it}$ ). The logarithm of the odds has the wanted property of not being restricted from 0 to 1, but extends from  $-\infty$  to  $+\infty$ .

In contrast to the probability model, the relationship between the independent variables and the log-odds of onset can viably be assumed linear. A simple multivariate logistic model can be written as:

$$L_{it} = \text{Log} (P_{it} / (1-P_{it})) = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + e_{it},$$

where  $\beta_0$  is the intercept,  $\beta_1, \beta_2, \dots, \beta_k$  are the coefficients to be estimated of the corresponding  $X_1, X_2, \dots, X_k$  independent variables, and  $e_{it}$  is the stochastic error term. Each coefficient ( $\beta$ ) measures the change in log-odds of onset for one unit increase in the corresponding independent variable ( $X$ ), holding all the other  $k$  independent variables constant.

## 4.2 The multicollinearity problem

Some of my independent variables are quite highly correlated. Hence, there is a potential collinearity problem related to the statistical analysis. This problem is largely inherent to my research question. I have intentionally selected independent variables that are hypothesized to be correlated with per capita income. If there were no multicollinearity when including all these variables in the same logit estimation, it would mean that the measures are not valid or all the explanations are completely wrong.

---

<sup>35</sup> The time period considered varies between the models, however, depending on how far back in time we have data for the explanatory variables. The civil war data are described in Section 5.1.2 below.

Multicollinearity concerns the degree to which one of the independent variables is a linear combination of the other independent variables in a model. Perfect multicollinearity means that  $X_k$  is a straight-line function of the other independent variables. High multicollinearity leads to estimation problems. To assess the degree of multicollinearity we can regress each independent variable on the other independent variables and find the  $R^2$ -values of each regression ( $R^2_1 X_1 | X_2, X_3, \dots, X_k$ ). A *variance inflation indicator* (VIF) is then usually calculated to see how the variance of an estimate is inflated by increasing multicollinearity:

$$VIF_1 = 1 / (1 - R^2_1)$$

The variance of the coefficient estimates is directly proportional to the VIFs, seen by:

$$\text{var}(\beta_1) = (\sigma^2 / \sum X^2_1) * VIF_1$$

A high variance of the estimators makes precise estimation difficult. This is reflected in higher standard errors and wider confidence intervals. In consequence, the risk of accepting a *false* null hypothesis of no relationship (type II error), increases. Hence, one or more coefficients of the highly correlated variables are often insignificant even though they may be quite large. The problem is quite similar to other data problems that reduce the possibility for making inferences, such as a small N and a low variance in the independent variables (Gujarati 2003:341-356). Consequently, high multicollinearity is particularly troubling in those of my models that have few observations. In these models, I try to assess how it may affect the results.

There are no good econometric solutions to the multicollinearity problem, as it essentially derives from properties of the data. Several remedies have been suggested, but these usually create greater problems than they solve. Ridge regression and principal components estimators are both attempts to reduce standard errors, but they do this at the cost of biasing the parameter estimates (Greene 1997:424-427). An obvious way to reduce the negative consequences of multicollinearity is to increase the number of observations. Several of my independent variables have many missing observations. For most of these variables I use linear interpolation to estimate missing values where possible.<sup>36</sup> Also important, I exclude all variables from the analysis that are not theoretically meaningful or that do not affect both the independent and the dependent variables. Control variables that do not satisfy these criteria

---

<sup>36</sup> See Section 5.3.2 for further explanation.

only reduce efficiency, and can be excluded without biasing the parameter estimates (King et al. 1994:183).

### **4. 3 The logic and structure of the empirical tests**

The research question is assessed by testing the individual hypotheses derived in the previous chapter. I test the hypotheses one after another rather than by including all the variables in a single model. The first reason is that testing all in a single regression would mean an insurmountable multicollinearity problem. The second reason is the need to maximize the number of observations. Some of the independent variables have many more missing observations than others. Including them in all models will thus weaken the tests of the other variables. The more variables included in one model, the smaller N. Thirdly, by testing the explanatory variables one by one I can more easily assess their individual relationship to *per capita income*, which is central to most of the hypotheses.

Except for the hypotheses of the greed explanation, all the hypotheses hold that some explanatory variable can account for parts of the *per capita income* effect. The explanatory variable is thought to be confounding, intervening, or interacting with *per capita income*, as shown by the causal diagrams in chapter three. However, it was not possible to specify precise causal models; in most cases the theory did not point out whether the explanatory variable was confounding or intervening. Most of the causal diagrams were used to illustrate the logic of the argument, not to derive completely specified, testable models. Consequently, I will not perform a path analysis in order to estimate the exact size of the direct, indirect, interaction, and spurious effects.

Event though most of the explanations do not allow me to specify exact causal paths, they provide testable implications. The central point is that the effect of *per capita income* on the dependent variable should be weakened both when controlling for an intervening and a confounding variable (Skog 2004:276). In most of the hypotheses I have stated that the *per capita income* effect should be *considerably* weakened when controlling for the variable in question. To systematize the testing I suggest a threshold for what is considerable, even though such a threshold must be arbitrary. I suggest that if the *per capita income* effect is not weakened by at least 20 percent, the null hypothesis of no relationship should be kept.

Unstandardized beta coefficients are used as effect estimates primarily because there is no consensus about how to estimate standardized coefficients in logistic regression (Menard 2004).

The second part of the hypotheses concerns the effect of the explanatory variable in question on the risk of civil war onset. Clearly, variables that have the opposite of the predicted effect can be dismissed. It is more difficult to determine when a coefficient having the correct sign is *significantly* related to the dependent variable. This is due to the problems of collinearity combined with a low number of observations. If the explanatory variable (X) can account for all of the *per capita income* effect, X should have a significant effect even when *per capita income* is controlled for, as Skog (2004:276) explains. However, when *per capita income* and X are highly correlated and N is small, the effect of X may not turn up as significant due to the difficulty of estimating its precise individual effect. Though X may be the causally most important variable of the two, it may still have a relatively high p-value.

Thus, I need to assess estimated significance levels with caution when it comes to highly correlated variables. One way of assessing this is to look at what happens with x when *per capita income* is excluded for the model. It should then become significant if the hypothesis is correct. If this significance is due to its standard error being much reduced, it can be taken as a sign of collinearity and imprecise estimation causing its non-significance in the joint model. If it is rather due to a large increase of its coefficient, it may indicate that *per capita income* accounts for much of its effect, contradicting the hypothesis. This is not entirely certain, however, as coefficients can also be affected by high levels of collinearity (Greene 1997:420). Thus, some degree of judgment must be used when assessing significance.

The hypothesis testing can help us both to find which lower-level accounts are the most plausible. I have discussed many different accounts (belonging to the four different types of explanations), and derived testable implications from them. It is very unlikely that they are all correct. An important task is to find out which of these receive no empirical support, and can be considered implausible. This may help us remove at least parts of the interpretative uncertainty surrounding the poverty-civil war relationship. Perhaps more importantly, the tests of the hypotheses can help us to evaluate the strength of the four categories of explanations. If none of the hypotheses derived from one explanation receive any empirical support, it can be considered weakened.

#### 4.4 Methodological problems and remedies

Apart from the data problems, there are at least four challenges facing analysts of quantitative civil war data: non-independent observations, unmeasured heterogeneity, endogeneity, and the rareness of onset events (Gates 2002:1). First, logistic regression assumes that observations are *independent* of each other. The risk of civil war breaking out in Chad in 1995 is assumed to be independent of what happened in Chad in preceding years. This assumption of temporal independence seems untenable. Having a civil war in the previous year plausibly makes a country more likely to have a new civil war onset in the next year. Researchers correct for this in different ways. I use Fearon and Laitin's (2003a) method of adding a dummy variable which captures whether the country had an ongoing war in the previous year or not. They call this variable "*prior war*", but I find "*ongoing war*" a more appropriate label. Somewhat surprisingly, their results show that *ongoing war* is negatively related to the risk of civil war onset (Fearon and Laitin 2003a:84).<sup>37</sup> Thus, having an ongoing civil war in one year decreases the chance of a new civil war breaking out in the next year. A possible explanation can be that most countries are too small for having two separate civil wars ongoing at the same time.

Other studies have found the expected positive relationship between having a past civil war and the risk of a new civil war onset (e.g., Urdal 2005:427). These studies do not look at whether there was an ongoing war in the prior year, however, but rather how much time has past since the last war ended ("*brevity of peace*"). This is often captured by a variable taking the value 1 immediately after a conflict ended, and decreasing at a certain rate for each following year (cf., Raknerud and Hegre 1997:393). The rationale is that the temporal dependence effect is strongest right after the last war ended, and then gradually decreases with time. In addition to the control for ongoing war, I add this variable, *brevity of peace*, using the formula  $\exp(-\text{years since the last war ended} / x)$ , where  $x$  determines the rate of decreasing effect.<sup>38</sup>

The problem of *spatial dependence* is less recognized in the quantitative conflict literature. Still, much evidence suggests that conditions in neighboring countries matter for the risk of

---

<sup>37</sup> Also surprisingly, they do not mention this counterintuitive finding.

civil war (Gleditsch 2007). To reduce this problem, I add a control variable, *neighborhood wars*, which measures how many civil wars were ongoing in contiguous countries in the previous year.<sup>39</sup> This should capture some of the potential spillover-effects that are likely to increase the risk of civil war onset in a neighboring country. I mark contiguous countries on the basis of Gleditsch and Ward's (2001) "Minimum distance data" (distance=0).

The problem of *unmeasured heterogeneity* is the most difficult to handle. Causal analysis of cross-national data rests on the quite untenable assumption that units are homogenous; that is, the expected value of the dependent variable should be the same for every unit when the explanatory variables takes on a particular value (King et al. 1994:91). It is not plausible that the risk of civil war is the exact same in too different countries if they have the same values on the explanatory variables in my models. This can also be understood as an omitted variables problem; we lack some variables that condition the effect of the explanatory variables in different contexts.

Attempts to alleviate the problem in civil war studies have been unsatisfactory. *Fixed effects estimations* attempt to control for heterogeneity by adding dummy variables for every country. The great problem of this method is that all countries displaying no variation in the dependent variable have no impact on coefficient estimates. In analyses of civil war this means that the majority of countries are excluded, as they did not experience civil war in the period under study. Collier and Hoeffler (2004:583-587) test whether fixed effects estimations affect the results of their model. They lose almost 80 percent of the observations by using this method, and the results are not surprisingly changed. I would thus argue that using fixed effects models adds an even greater methodological problem of few observations. The heterogeneity problem could also be lessened by focusing on a specific region or cluster of relatively similar countries. This is not a good option for answering my research questions, however, which involve the difference between rich and poor countries. All the existing variation in the explanatory variables should be included in the analysis to maximize the

---

<sup>38</sup> I test different values of x to see which gives the model with the best fit to the data (see section 6.1).

<sup>39</sup> It looks at the previous year to reduce endogeneity. For the first year in each country's time-series, I use data of neighborhood wars for the *same* year to avoid losing several observations of civil war onsets. This is a somewhat problematic procedure here, but since the variable is not theoretically central to my study, I think the risk is not large. I also test this variable as a dummy, taking the value 1 if there are one or more wars in contiguous countries, and 0 if no war. I find the first operationalization more appropriate, however, as it captures the extra risk of spill-over effects from *more* than one war in the neighborhood.

possibility for making inferences. Thus, like most quantitative civil war studies, I choose the risk of bias due to unmeasured heterogeneity over the alternatives.

*Endogeneity* problems are significant in the study of civil war onset. Using a single-equation method, we assume that the causal effect goes only from the independent variables to the dependent variable. We know, however, that civil war affects the values of some of the central independent variables (cf. Thies 2004). The usual way of controlling for endogeneity is to lag the explanatory variables one year (Sambanis 2002:238). This serves to strengthen the assumption that the causal sequence follows the theoretical model; the values of the explanatory variables are assigned before the value on the dependent variable. I do this for all variables with plausible endogeneity problems. In one analysis I also try a different way of reducing endogeneity. I include a control variable which may remove parts of the plausible endogeneity from the causal variable. The logic of this method is more thoroughly explained in Section 6.2.4.

Finally, the *rareness of civil war onset events* is a challenge for the analysis. Qualitative conflict researchers are often charged with selection bias due to their selecting cases on the dependent variable. Quantitative researchers, though avoiding the problem of selection bias, face another problem by including a vast amount of “non-onsets”, and comparatively few onsets.<sup>40</sup> Using standard logistic regression, this causes some amount of bias in coefficient estimates and underestimation of the probability for onset, particularly when the total number of observations is not great (King and Zeng 2001a:146).<sup>41</sup>

King and Zeng (2001a) have developed a “rare events logit” method which attempts to correct for this bias. I tested the method on my models. It did reduce the standard error of some coefficients in the models with relatively few onsets (below ca. 70). However, a large disadvantage is that log likelihood values are not displayed when using their rare events software. These are central to my analysis, particularly because several of my variables are highly correlated. Second, some very implausible results appeared in some models which

---

<sup>40</sup> In our (the Fearon-Laitin) dataset, civil war onsets make up 1,72 percent of all country-years. The rare events problem is alleviated in the PRIO/Uppsala armed conflicts dataset, which uses a lower death threshold to code a conflict. We do not choose the latter as our primary dataset due to theoretical considerations, which we discuss below.

<sup>41</sup> See King and Zeng (2001a:145-150) for a more extensive explanation.



included variables with large dispersion and highly correlated variables.<sup>42</sup> I therefore choose to use normal logit estimation and interpret the results with caution in the models with few onsets.

#### 4. 5 Choosing a set of control variables

I do not control for all the variables included in Fearon and Laitin's (2003a) model. Many of their variables can be excluded because they are not significantly related to the dependent variable (see Section 6.1). One of the significant variables from their model, *mountainous terrain*, can be excluded because it is not plausibly related to per capita income or the other independent variables. I also exclude the *new state* variable, though it is robustly related to civil war onset. Controlling for this variable would probably confuse my analysis. Fearon and Laitin (2003a:81) argue that it is a proxy for state weakness, but it can plausibly be related to other theoretical variables (see Section 3.2.5). This makes sensible analysis of my state capacity proxies difficult, as it will share some of their variation, and some of other relevant theoretical variables. *Instability*, which denotes whether the country had a large change of regime type during the three years before the observation, is also excluded from my models. This is primarily due to its potentially large endogeneity problem. It might very well be that violent conflict causes political instability, and not the other way around.

I include *oil exporter* from the Fearon-Laitin model as a control variable because it is plausibly related to both *per capita income* and the dependent variable. Controlling for *oil exporter*, the *per capita income* effect may perhaps be more clearly interpreted as a measure of economic development. Many oil states have less diversified, high-tech, and productive economies than their *per capita income* might suggest. Thus, controlling for *oil exporter* can make the *per capita income* variable more theoretically meaningful. I also include *anocracy*, which marks countries that mix democratic and autocratic characteristics, since it is probably related to both *per capita income* and the risk of civil war onset. Hegre (2003) finds evidence for an interaction effect between level of development and type of political regime.

*Population*, a standard control variable in the civil war literature, is also included (Hegre and Sambanis 2006:514-515). This variable is important because it strengthens the assumption of

---

<sup>42</sup> It is difficult to assess why these peculiar results appeared. I can only speculate that the rare events logit seems

homogenous units. Without controlling for population, we assume that the chances for civil war in Gambia should be equal to that in China if they had the same values on our explanatory variables. Clearly, *ceteris paribus*, an armed conflict is more likely to reach the threshold of 1,000 deaths in more populous countries. Controlling for population thus reduces the unmeasured heterogeneity problem. Finally, in all the models I include the two controls for temporal dependency described above, and *per capita income*, which is at the centre of the analysis. This amounts to a total of six control variables.

## 5 Indicators and data

In this chapter I account for the measures of the variables described in Chapter three. Most of the variables have to some extent been operationalized in the theoretical analysis. Finding more concrete indicators of abstract theoretical concepts, such as economic opportunity, was essential to the derivation of testable hypotheses. Thus, for most of the variables I here account for their exact measurement, the data that are used and possible transformations. I also consider reliability, but the discussion of possible bias of results is found in the empirical analysis. I begin with the dependent variable, proceed to the explanatory variables, and finally consider the control variables. Descriptive statistics for the variables is found in Appendix 1.

### 5.1 Civil war onset

#### 5.1.1 Choosing a civil war measure

Operationalizing civil war onset is a complex task. Several civil war lists exist that differ in their definitions of civil war and in their coding rules. I consider it expedient to apply these existing lists, and not engage in my own coding, as this is time-consuming and demanding work. In choosing between these, I first consider content validity; the degree to which the measure reflects my theoretical definition. It should only include conflicts where (a) fighting is internal to the state, (b) the government is a belligerent party, (c) both sides put up effective resistance, and (d) fighting causes a relatively high number of deaths. Several civil war datasets largely keep with this general definition. I considered three of these which are all much used in the literature.

The Correlates of War (COW) project has for decades been the leading supplier of conflict data, and most other datasets draw heavily on this work. Although the COW data have been invaluable for quantitative conflict studies, much confusion surrounds how these data are coded. This owes partly to lack of transparency, and partly to coding rules being changed at least three times over the years (Sambanis 2004b:817). Most importantly, in the early COW data the death threshold was 1,000 a year, but this was (apparently) changed in 1992 to a cumulative threshold of 1,000 deaths during the course of the war (Gleditsch et al. 2002:617). Many researchers still believe that the old definition applies. Sambanis (2004b:817) also points out that it is unclear how they implicated these new coding criteria backward in time.

The lack of transparency is the main reason why I do not choose COW (2000) as my primary dataset.<sup>43</sup>

The PRIO/Uppsala armed conflict datasets have also been much used in recent years (Gleditsch et al. 2002). They distinguish between (intrastate) minor armed conflict, intermediate armed conflict, and war. *Minor armed conflict* has a threshold of at least 25 battle-related deaths per year and less than 1,000 deaths accumulated over the entire war. I concur that this threshold is too low to count as a civil war.<sup>44</sup> *Intermediate armed conflict* produces at least 25 and less than 1,000 battle-related deaths per year, and at least 1,000 during its total course. I also think that this threshold is too low to qualify as a civil war. Finally, intrastate *war* has a threshold of 1,000 battle-related deaths per year. I think this operationalization better reflects my conception of civil war, but that its threshold may be too high. It excludes for example the Northern Ireland conflict which broke out in 1969, produced more than 25 deaths annually from 1971-1993, and claimed more than 3,000 lives in total (Gleditsch et al. 2002:617).

There are two other reasons why I do not use any of the PRIO/Uppsala datasets as a baseline. First, they do not include a threshold of deaths on both sides which ensures that effective resistance has been made. Second, they do not consider civilian deaths. This operationalization thus fails to capture that in most civil wars “the battlefield is society itself” (Kalyvas 2006:54). Most of the deaths in civil wars are civilian. Theoretical and empirical studies have concluded that this violence is often not purely criminal or opportunistic, but has a strategic, military aim (Azam and Hoeffler 2002:461; Kalyvas 2006). As civilian deaths are integral to the logic of violence in civil war, they should be counted into the threshold.

I choose to apply Fearon and Laitin’s (2003a) list of civil wars. The conflicts they include met the following criteria:

“(1) They involved fighting between agents of (or claimants to) a state and organized, nonstate groups who sought either to take control of a government, to take power in a region, or to use violence to change government policies. (2) The conflict killed at least

---

<sup>43</sup> Another reason is that updates only reach 1997, two years less than Fearon and Laitin’s (2003) data.

<sup>44</sup> We will, however, investigate whether some of the theories may apply also to minor armed conflicts.

*1,000 over its course, with a yearly average of at least 100. (3) At least 100 were killed on both sides (including civilians attacked by rebels).” (Fearon and Laitin 2003a:76)*

This operationalization has several advantages. First, I think the relatively high cumulative threshold of violence (1,000) is reasonable. At the same time, it allows for violence levels to fluctuate over time by accepting an *average* of 100 deaths per year. Second, it ensures that both belligerent parties put up effective resistance by the criterion of at least 100 deaths on both sides. Third, it includes civilian casualties. Fourth, and perhaps most importantly, they make explicit what rules they apply to handle several difficult coding issues. For example, civil war onset is coded in “the first year in which 100 were killed or in which a violent event occurred that was followed by a sequence of actions that came to satisfy the primary criteria” (Fearon and Laitin 2003a:76).<sup>45</sup>

Fearon and Laitin’s (2003a) operationalization differ from the other datasets particularly on one point: they see no theoretical reason for excluding anti-colonial wars within empires from their civil war list. They argue that to exclude such wars would be equal to excluding the Chechenya conflict in Russia if the Chechens finally should win independence. The authors admit, however, that there are large practical and theoretical problems with including the old anti-colonial wars. I do not have reliable data for most of the independent variables of interest, and it is obviously difficult to estimate an average of for example GDP for the entire British empire. They therefore have one civil war list including anti-colonial wars, and one excluding them. I will (as they primarily do) use the one without anti-colonial wars.

### **5.1.2 The Fearon-Laitin civil war data**

The Fearon-Laitin dataset includes all the 161 countries with populations above 500,000 for the period 1945-99.<sup>46</sup> The data are structured according to *country-years*; it includes all available years of all countries, and makes the country-year (e.g., France, 1950) the unit of analysis. Civil war onset is a dummy variable, coded “1” for country-years with onset and “0” for country-years without onsets. Country-years with ongoing war are also coded “0”, and new onsets are coded “1” even if there is already an ongoing war in the country. This coding practice is disputed, as it may increase the problem of dependency between observations. It is

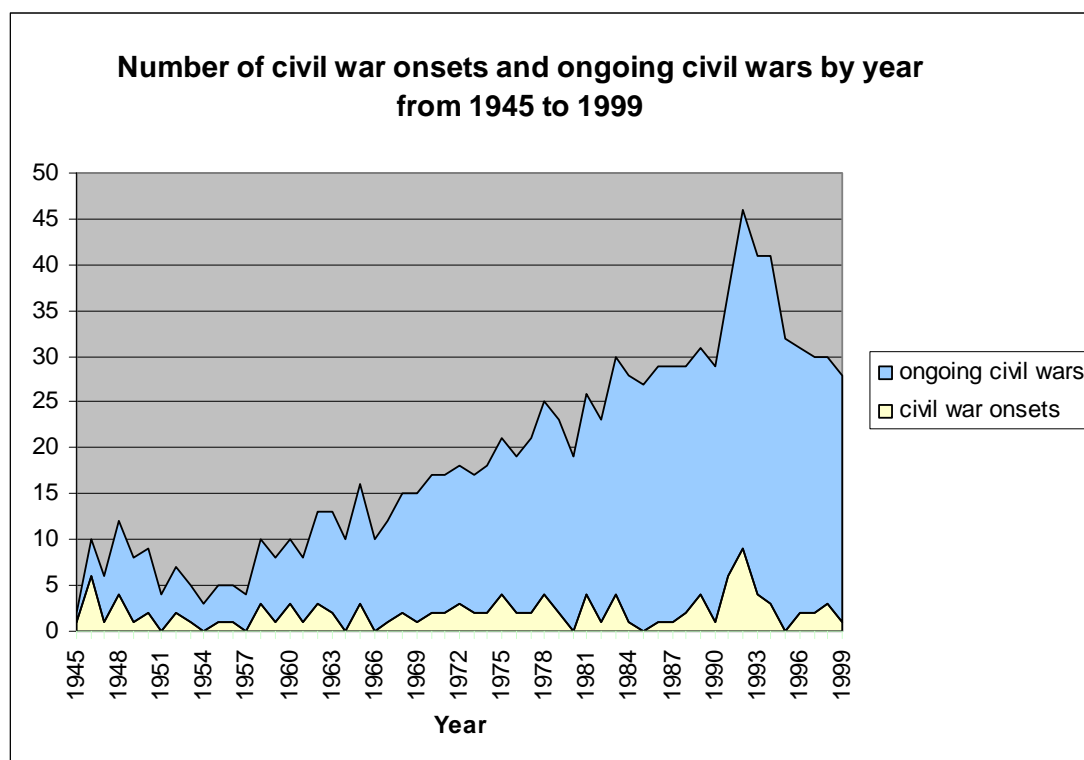
---

<sup>45</sup> See the other coding rules in a footnote in Fearon and Laitin (2003:76).

plausible that an ongoing civil war makes a country more prone to another rebellion breaking out. On the other hand, ignoring all new onsets in countries with ongoing civil wars may present a larger problem. In this dataset, that would mean dropping 14 civil war onsets, thus losing valuable information and significantly increasing the problems of very rare onset events.

Civil war onset is a rare event. There were a total of 114 civil war onsets in the 1945-99 period.<sup>47</sup> Onsets thus make up 1,72 percent of all 6625 country-years. Most civil wars broke out in sub-Saharan Africa (34), Asia (33), and North Africa/Middle East (17), followed by Latin America (15), Eastern Europe and the former Soviet Union (13), and what Fearon and Laitin names “the West” (2) (Western Europe, USA, Canada, Japan, Australia, and New Zealand).

**Figure 9.**



<sup>46</sup> They include country-years below the 500,000 population threshold if the country reaches the 500,000 inhabitants threshold later in the series. No civil war onsets occurred in country-years with fewer than 500,000 inhabitants. The closest is Djibouti, which had 565,000 inhabitants when the civil war started in 1993.

<sup>47</sup> See the war list in Appendix 2(?)

Figure 9 shows that the rate of aggregate civil war onsets per year has been relatively stable over time. On average, 2.07 civil wars have started each year. The largest divergence from this trend is the turbulent post-Cold War years of 1992 and 1993, where respectively 6 and 9 civil wars started. The number of *ongoing wars* in the aggregate has on the other hand increased sharply to a peak of 38 wars in 1994, after which it declined somewhat. This is not primarily due to an increase in the rate of onsets, but rather due to an accumulation of enduring wars, as Fearon and Laitin (2003a:77) points out.

## 5.2 Income per capita

I consider per capita income primarily a measure of economic development, and use it as an indicator of economic opportunity (Section 3.4.3). However, I describe the operationalization of per capita income before the other variables because it is central to nearly all the analyses. I apply Fearon and Laitin's (2003a) GDP per capita estimate. This is based on Penn World Tables (PWT) 5.6 data for "RealGDP", measured in constant 1985 U.S. dollars. Those data cover the 1950 to 1992 period.<sup>48</sup> Fearon and Laitin extend the series to 1999 using the World Bank's (2001) *World Development Indicators* (WDI) growth rates. Then they estimate missing values using data on energy consumption from the Correlates of War Project (COW). These methods produce a very complete series, covering 6373 observations, and all of the 110 onset events in the dataset. The values are lagged 1 year to alleviate the endogeneity problem of civil war affecting economic productivity (Fearon and Laitin 2003b:1-3).<sup>49</sup> Unlike Fearon and Laitin, I log-transform the variable. They report that log-transformation had no impact on the results (Fearon and Laitin 2003a:87). I find that it *does* matter in tests where fewer observations are included. Log-transformation is recommendable because it is likely that one unit change has greater effect for countries with low values than for countries with high values on this variable. It also reduces the variable's large positive skewness.

---

<sup>48</sup> Their methods of extending the data and filling in missing values are complicated, and are only briefly described here. For the entire procedure, see Fearon and Laitin (2003b:1-3).

<sup>49</sup> The first observation in each country's time-series is not lagged, as this would mean losing 10 onset events. They argue that using the contemporaneous estimate for this first observation for each country should not be very problematic because the values do not change very much from year to year. We concur, and use the same practice for other variables, such as the tax ratio.

## 5.3 Grievance indicators

### 5.3.1 Extraction level and change in extraction (H1 and H2)

Level of extraction is commonly measured by *the tax ratio* – the state’s tax revenue as a percentage of GDP (e.g., Fauvelle-Aymar 1999; Chaudhry 1997). The World Bank (2003) defines tax revenue as compulsory transfers to the central government for public purposes. Non-tax revenues and certain compulsory transfers such as fines, penalties, and social security contributions are excluded. The World Bank used to collect most of these data until 1975, when the International Monetary Fund (IMF) assumed responsibility for the collection. According to Johnson and Rabinowitz (2005:12) these institutions “share some standardization in reporting format and methods”. There are still considerable problems related to using data collected from the different institutions, which is necessary to get a comprehensive dataset. Johnson and Rabinowitz (2005) have taken on the task of reconciling data from the different institutions, and also filling in missing data from central bank reports. I therefore use their data, which are the most comprehensive and probably most reliable data available.<sup>50</sup> Since their data set does not include all the countries of the Fearon-Laitin dataset, I supplement with data from the World Bank (2006).<sup>51</sup>

I measure *change in extraction levels* by the percentage change in tax ratio from 3 years before to 1 year before the observation. This operationalization tries to balance the need to minimize the potential endogeneity problem and the problem of few onset events. The likelihood of bias due to endogeneity could perhaps have been more reduced by looking at change from further back in time. I choose not to do this since it would mean losing observations for the early years of each country’s time-series. This is unfortunate especially because there are several onset events in these years.<sup>52</sup> The variable has a very large dispersion (min = -93.33 and max = 600). Particularly, some extreme outliers in the top range make the standard deviation very large. I choose to exclude values above 200 since I consider it very unlikely that any country has more than 200 percent increase of tax ratio in a two-year period. This means dropping 24 country-year observations and no onset events. With my operationalization, 3799 observations, 125 countries and 61 onset events are included.

---

<sup>50</sup> Still, reliability is probably not very high. We discuss this in chapter 6.

<sup>51</sup> The supplement amounts to 115 country-year observations.

<sup>52</sup> For example, looking at tax ratio change from 5 years before to 1 year before the observation, we lose 5 onset observations from the analysis.



### 5.3.2 Public goods provision (H3)

I use *gross total primary school enrollment rates* as an indicator for public goods provision. Providing primary education for all is widely seen as a basic obligation of the state. This is seen at the international level by the fact that most states have obliged themselves to make primary education available free for all through the UN Convention on the rights of the child (United Nations 1989:8). It is difficult to tell whether people of very poor societies expect the state to arrange for education, and how important this is to them. Some authors suggest that young people also in poor African countries have come to expect this, and to value education highly. Peters et al. (2003:25), for example, hold that grievances and frustration related to the collapse of educational opportunity structures in the 1980s and 1990s in Sierra Leone was an essential motivating factor for the RUF rebels.

The provision of education could also be measured by education expenditure. Enrollment rates are preferable for at least three reasons. First, expenditure data may reflect spending on universities as much as primary schools. Most likely, the lack of tertiary education is not a major grievance for most people in poor countries. Second, primary school enrollment is an outcome variable, which captures what education services people are actually offered. This also increases reliability. It is easier to count how many pupils are in school than to check government spending. Third, the data coverage is better for enrollment rates than for education expenditures.

The data on gross primary school enrollment come from the UNESCO Institute for Statistics (2000), and covers the period 1970-1997. There were many missing observations in the raw variable. I increased the coverage by linear interpolation, which is commonly used for education variables in the literature (e.g., Fearon and Laitin 2003b:23).<sup>53</sup> Interpolation estimates missing data points between two known data points on the basis of the assumption that there is a constant rate of change between these two known observations (Spalter 2007). This assumption should not be violated here, as the rate of change in primary schooling seems very stable. The variable was then lagged one year.

---

<sup>53</sup> Other studies in addition use more controversial imputation methods. Thyne (2004:739) uses “Amelia imputations”, and Fearon and Laitin (2003b:23) extrapolate.

## 5.4 State capacity indicators

In section 3.2 I suggested substituting Fearon and Laitin's (2003a) very crude state capacity proxies with four proxies that can be more easily interpreted. One of these, the *tax ratio*, is already described in section 5.3.1. I account for the three other variables below.

### 5.4.1 Counterinsurgent capacity (H4.1)

I have argued that the number of military personnel *per capita* is probably the best available cross-national indicator of counterinsurgent capacity (section 3.2.6). Data on military personnel is preferable to data on military expenditures, which are well known to be extremely unreliable (Herbst 2004:364). I get data for military personnel from the Correlates of War Project (COW), which is the most complete dataset. They seek to measure the capacity to fight *interstate* wars, not domestic insurgents. Most irregular forces, which the literature on counterinsurgency points out as essential, are thus excluded (Correlates of War 2005:8). Despite its shortcomings, I stick to the argument that *military personnel* is a better proxy for counterinsurgent capacity than *per capita income*. The variable is specified as military personnel per 1000 inhabitants. I used logarithmic transformation to reduce its heavily skewed distribution and lagged it one year to reduce endogeneity. The data coverage for the variable is very good.

### 5.4.2 Road density (H4.2)

I have argued that *road density* can be used as a proxy for the state's reach into rural society (section 3.2.7). I use Canning's (1998) data on the length of *paved roads* per square km to measure this. Paved roads are defined as concrete or bitumen-surfaced roads (Canning 1998:535).<sup>54</sup> This is to my knowledge the most complete existing series of road stocks, as it includes data from the International Road Federation (IRF), the UN, as well as national sources (Canning 1998:534). It has relatively good coverage for 152 countries from 1950 to 1997. I complemented the Canning (1998) data with World Bank (2006) data, extending the series to 1999, and filling in missing observations where possible. This amounted to 712 additional observations. After interpolating these data, there were 4512 observations, including 65 onset events.

---

<sup>54</sup> We choose not to use a measure for total roads, which includes unpaved roads, because these data are much less reliable (Canning 1998:531-536).

### 5.4.3 Government share of consumption (H4.3 and H5)

Government share of consumption was argued to be a proxy for the state's available resources, or economic muscle (Section 3.2.8). Finding the most valid and reliable measure of this variable proved to be difficult. There is great variation in this measure derived from two commonly used sources: the WDI 2003 (World Bank 2003) and the PWT 6.2 (Heston et al. 2006). The WDI 2003 and PWT 6.2 variables, which should presumably measure the same concept, correlate only at  $r = 0,40$  in my dataset. This could be due to different definitions and standards of measurement, or low reliability of the data from one or both sources. The WDI 2003 definition includes all government current expenditures for purchases of goods and services - also payment of public sector employees. Expenditures on national defense and security are included, except military expenditures that are part of government capital formation (World Bank 2003). Unfortunately, there is little available information concerning the definition and sources of the PWT 6.2 data. Heston et al. (2004:3) state that they construct the "Real GDP" measure by combining its three component shares: private consumption, government consumption, and investment. They do not explicitly report how they distinguish between these three component parts.

The publishers of the PWT 6.2 argued that the discrepancy between the WDI 2003 and the PWT 6.2 data is due to the PWT data being purchasing power parity (PPP) adjusted, whereas the WDI data are in national currencies.<sup>55</sup> This clearly affects the estimation of GDP, but I cannot see that it significantly affects the *government share* of GDP. The World Bank gave a more probable explanation: whereas the WDI 2003 data rely on the definition of the IMF's Government Finance Statistics (GFS) manual of 1986, the PWT 6.2 might be using data corresponding to the IMF's later GFS 2001 manual.<sup>56</sup> These differ in two ways. First, the accounting method of the GFS 1986 version is cash-based, whereas the GFS 2001 version uses an accrual accounting method. The latter considers not only cash transactions, but also depreciation of fixed capital. Second, the GFS 1986 version reports only current expenditures, whereas the GFS 2001 version also includes expenditure on capital formation (World Bank 2006: Table 4.10, "About the data"; IMF 2001:157-158).

---

<sup>55</sup> E-mail exchange between myself and Ye Wang, University of Pennsylvania, 08.10.2007.

<sup>56</sup> E-mail exchange between myself and Beatriz Prieto-Oramas, Development Data Group, The World Bank, 05.10.2007.

The GFS 1986 version used in the WDI 2003 is probably the most appropriate for my purpose. The current cash expenditures of the government seem more theoretically salient than capital expenditures. What happens to the government's capital stock is probably less relevant for the government's ability to prevent rebellion than the government's ability to currently spend on administration, police and military. The data available from the World Bank (2003) are also much more complete using the old GFS 1986 version. I do not know whether the PWT 6.2 variable rely on the IMF's GFS 2001 data collection. This lack of transparent reporting is another good reason for using the WDI 2003 data. I further discuss the clearly significant reliability problems in the next chapter.

Hypothesis 6 concerns decreasing revenues in countries with a weak state. I use *government share of consumption* as an indicator of government revenues. Consumption data is chosen because the standard revenue data from the World Bank and the IMF exclude grants and aid. Revenue data therefore do not capture the large decrease in revenues for many poor client states after the Cold War. This may have large consequences for my results as several of these countries experienced civil war in the 1990s. I use government share of (total) consumption instead of absolute values of government consumption because this allows us to control for changes affecting the whole economy, and not only government revenues. As for tax ratio change, I measure change in government consumption from 3 years to 1 year before the observation. I found that this specification best balances the need to control for endogeneity (assuring that the revenue change comes before the civil war), and to avoid many missing observations. The variable is operationalized as a decreasing revenues dummy (Section 6.3.4)

## **5.5 “Ecology” indicators**

In the theoretical analysis (section 3.3) I discussed three “ecological variables” that should affect the viability of insurgency: settlement structure, road density, and telecommunication infrastructure. The road density measure is described above (section 5.4.2). I here consider the other two measures.

### **5.5.1 Urban settlement (H6)**

I use data for *the percentage of the population living in urban areas* from the World Bank (2006). They use data reported from every country to the UN. Unfortunately they do not state

the degree to which standardized coding criteria have been used. Kocher (2004:58) reports that he investigated this by browsing through the various countries' census classifications. He finds that on average a settlement is counted as urban if it has more than 2000 inhabitants. The population threshold varies somewhat between countries, however. He did not do any systematic analysis of how much this affects reliability. The variable has good coverage from 1960 to 1999.<sup>57</sup>

### **5.5.2 Telecommunication infrastructure (H7)**

I use the *number of telephone sets per 1000 inhabitants* as an indicator for telecommunication infrastructure. It would be more theoretically appropriate to use telephone mainlines as an indicator. The reason I do not use this is that its data coverage is lower than for telephones, and the two measures are almost perfectly correlated ( $r = 0.99$ ). The data for telephones extend from 1950 to 1995, with reasonably good coverage. I linearly interpolate also this variable, giving 117 more observations. Although this variable is more stable than *GDP per capita*, there might be some risk of endogeneity to civil war. I lag all the values of the variable one year, except the first in each country's time-series (as for *GDP per capita*), to reduce the potential endogeneity problem. The variable is also log-transformed to reduce skewness, and because it is plausible that an increase at lower values of the variable has a stronger effect than at higher values.

## **5.6 Economic opportunity indicators (greed explanation)**

In the theory chapter (Section 3.4.3) I suggested five indicators of economic opportunity in the regular labor market, which can be used to test the greed explanation: *per capita* income, economic growth, male secondary schooling, income inequality, and level of extraction. The income *per capita* measure is already described (Section 5.2). The measure used for level of extraction is the *tax ratio*, which is described above (Section 5.3.1).

### **5.6.1 Economic growth (H8.2)**

There is large variation in available estimates of economic growth. One reason is that some measure it in local current currency (World Bank 2006), while others use standardized estimates in constant US dollars (Heston et al. 2006). I follow the latter standard, as for

---

<sup>57</sup> Kocher (2004) imputes missing values from the COW urbanization variable. We do not think this is a very

income *per capita*, and use the annual economic growth rate that is estimated in the PWT 6.2 dataset. This is based on their constant dollar “Real GDP” chain variable (Heston et al. 2004:11-12).<sup>58</sup> The variable is lagged one year to reduce the problem of endogeneity. This probably does not eliminate the endogeneity issue, which is particularly troubling for this variable. We consider possible bias in the empirical analysis.

### **5.6.2 Male secondary schooling (H8.3)**

Data for gross male secondary school enrollment rates were found at the UNESCO Institute for Statistics (2000) and cover the years 1970-1997.<sup>59</sup> As for primary school enrollment, I linearly interpolated to decrease the number of missing observations. The assumption of constant rate of change should not be severely violated for this variable either.

### **5.6.3 Income inequality (H8.4)**

I use data for income inequality from the UNU/WIDER *World Income Inequality Database 2b* (UNU/WIDER 2007). They measure income inequality by a GINI coefficient. I have only included GINI observations that were estimated for a large part of the country (not for example only for rural or urban areas). Most of the data extend from 1960-1999, but a few countries have observations before 1960. The data are generally quite scarce, and particularly so for developing countries. I use linear interpolation to increase the data points. This is not problematic, as income inequality does not fluctuate much over time. Still, there is a problem of low coverage. There are inequality data for 3322 observations, 137 countries, and for only 44 civil war onsets.

## **5.7 Control Variables**

In section 4.4 I discussed which control variables to include, and ended up with five variables besides *per capita* income. The controls for temporal dependency - *ongoing war* and *brevity of peace* - were sufficiently described in section 4.4.2. I here shortly describe the three remaining control variables - *population*, *oil exporter*, and *anocracy* - which are all taken from the Fearon-Laitin dataset.

---

reliable method, as COW use different definitions of urban population.

<sup>58</sup> We consider this a safer option than using the Fearon-Laitin GDP *per capita* estimate to calculate growth rates. They have extrapolated and used various imputing methods which may affect the growth rates.

<sup>59</sup> We cannot extend the series to 1999 because UNESCO changed their definition and standards of collection after this year.

The *population* variable is based on World Bank data. The variable is lagged one year to reduce endogeneity, and log-transformed to alleviate the problems of a skewed distribution (Fearon and Laitin 2003a:81).<sup>60</sup> *Oil exporter* marks countries in which fuel exports form more than one-third of total export revenues. It is based on World Bank data covering five-year intervals beginning in 1960. Fearon and Laitin (2003a:81) report that they have interpolated for years after 1960 and used the value for 1960 for years prior to 1960. They have also used some country-specific sources for a few countries without World Bank coverage. *Anocracy*, a dummy variable marking countries that are neither full-blown democracies nor autocracies, is based on the Polity IV data. Countries with values between -5 and 5 on the Polity IV regime scale (where -10 is most autocratic and 10 is most democratic), are assigned the value 1 on *anocracy*.<sup>61</sup> (Fearon and Laitin 2003a:81).

---

<sup>60</sup> For the raw population variable that is not lagged, we needed to add data for 1999 from World Bank (2006). This was useful when constructing the measure *military personnel per 1000 inhabitants*.

<sup>61</sup> This variable may have an endogeneity problem. The Polity III data considered violent political competition a semi-democratic trait, which clearly might make it endogenous to civil war. Strand (2007:5) reports that the issue was supposed to be addressed in the Polity IV update, but it is not clear how well it was done. This potential endogeneity should have no major consequences for my analysis, however, as I only use it as a control variable.

## 6 Empirical analysis

In this chapter I test the hypotheses derived in Chapter 3 using logit analysis of cross-national time-series data. I begin by replicating Fearon and Laitin's (2003a) model of explanatory variables, and then reduce it to the control variables suggested in Section 4.5. This constitutes the basic model which is used for testing all the individual hypotheses. The tests follow a similar logic, except for the greed hypotheses, which are somewhat differently formulated (see Section 3.4.3 and Section 4.3). I group the hypotheses by the explanation they are derived from. As some hypotheses are related to more than one explanation, this structure can not be strictly followed. Some subsections will discuss results that are also relevant to other explanations. The hypotheses are tested one after another, as explained in Section 4.3. I report previous findings for hypotheses that have been tested before in large-N quantitative studies before proceeding to my own empirical analysis.

### 6.1 From the Fearon-Laitin model to a new basic model

Fearon and Laitin's (2003a:84) third logistic regression model is replicated in Table 2, Model 1 (Model 2-1).<sup>62</sup> It shows that *income per capita* has a negative and highly significant effect when controlling for a set of political, geographic, and social variables.<sup>63</sup> The effect is also very strong in substantive terms. This can be gauged by looking at the odds ratio.<sup>64</sup> A coefficient of -0.32 corresponds to an odds ratio of 0.73. Thus, 1000\$ higher per capita income is associated with on average 27 % lower annual odds of civil war onset.

I have argued that *ongoing war*, *population*, *oil exporter*, and *anocracy* should be kept as control variables in my analyses (Section 4.5). Some of the other variables are irrelevant because they are not significantly related to the dependent variable. This is the case for *noncontiguous state*, the *democracy dummy*, *ethnic fractionalization*, and *religious fractionalization*, which do not reach significance even at the 0.10 significance level ( $\alpha$ ). All the control variables I suggest to keep are significant at  $\alpha = 0.05$  or lower. The *population* coefficient has the expected positive sign, implying that a larger population is related to a higher risk of civil war onset. As I have discussed, *ongoing war* is negatively related to the

---

<sup>62</sup> Their third model is replicated instead of the first because it includes the *anocracy* variable.

<sup>63</sup> From here I consequently write variable names in italics.

<sup>64</sup> The odds ratio here refers to the relationship between the annual odds of civil war onset when the independent variable takes the value 1 and when it takes 0. Odds ratio = Odds(x+1) / Odds(x). The relationship between the coefficient (b) and the odds ratio is: Odds ratio = exp(b).



risk of civil war onset. Having an ongoing war in the previous year makes a new civil war onset in the present year less likely (see Section 4.4). *Oil exporter* has a positive effect, implying that countries which receive much of their income from fuel exports have a higher risk of civil war onset. As found in several studies, *anocracy* has a positive effect (e.g. Hegre et al. 2001). Countries mixing democratic and autocratic traits seem to be more prone to civil war.

**Table 2.** Replication and changes to Fearon and Latin's (2003a) logit model of determinants of civil war onset.

Variable	Model 2-1	Model 2-2	Model 2-3	Model 2-4	Model 2-5
Income per capita <sub>a, b</sub> (In Model 3-5: logged values)	-0.32*** (0.071)	-0.34*** (0.068)	-0.71*** (0.11)	-0.69*** (0.11)	-0.69*** (0.11)
Population <sub>b</sub>	0.27*** (0.074)	0.31*** (0.066)	0.30*** (0.066)	0.28*** (0.069)	0.26*** (0.074)
Ongoing war <sub>a</sub>	-0.92*** (0.31)	-0.79*** (0.30)	-0.82*** (0.30)	-0.83*** (0.30)	-0.85*** (0.30)
Oil exporter <sub>c</sub>	0.75*** (0.28)	0.73*** (0.26)	0.78*** (0.27)	0.72*** (0.27)	0.72*** (0.27)
Anocracy <sub>a, c</sub>	0.52** (0.24)	0.73*** (0.20)	0.81*** (0.20)	0.80*** (0.20)	0.81*** (0.20)
Brevity of peace				0.31 (0.24)	0.29 (0.24)
Neighborhood wars					0.051 (0.096)
Democracy <sub>a, c</sub>	0.13 (0.30)				
%Mountainous terrain <sub>b</sub>	0.20** (0.08)				
Noncontiguous state <sub>c</sub>	0.43 (0.27)				
Instability <sub>a</sub>	0.51** (0.24)				
New State <sub>c</sub>	1.66*** (0.34)				
Ethnic fractionalization	0.16 (0.37)				
Religious fractionalization	0.33 (0.51)				
Constant	-7.02*** (0.75)	-6.43*** (0.65)	-6.85*** (0.64)	-6.68*** (0.65)	-6.59*** (0.67)
Pseudo R <sup>2</sup>	0.11	0.082	0.082	0.084	0.085
Log likelihood	-478.67	-494.42	-494.35	-493.52	-492.23
Observations	6327	6327	6327	6327	6293
Onset Events	109	109	109	109	109

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous.

In Model 2-2 I keep only *per capita income* and the four suggested control variables. The *income per capita* coefficient is not appreciably altered. Thus, excluding the other control variables did not change the result most central to my analysis. All the control variables are now significant at  $\alpha = 0.01$ . In Section 5.2 I argued that it is preferable to log-transform the

*income per capita* measure. This is done in Model 2-3. There are no large changes in the results (except that the *per capita income* coefficient must be differently interpreted as its scale is changed).<sup>65</sup>

In Model 2-4 I include the second control variable for temporal dependency, *brevity of peace*. As explained in Section 4.4, this variable takes the value 1 in the year after the previous civil war ended, and then decreases toward 0 as time passes. I use the function  $\exp(-\text{years since the last war ended} / x)$  to construct the variable, where  $x$  determines the rate with which the value decreases. I test for different values of  $x$  to find the decreasing rate giving the best fit to the data. I find that a very slow rate gives the model with the best fit measured by log likelihood. I settle with  $x = 400$ , implying a half-life of 277 years.<sup>66</sup> For slower rates of decrease, the fit with the model is not appreciably changed. The *brevity of peace* coefficient has the predicted positive sign, implying that the risk of civil war is higher the more proximate is the prior war experience. The coefficient is not significant at the 0.10 level. I still keep it as a control variable, because it becomes significant in some of the other models, and it does not cause much inefficiency.

In Model 2-5 I add the last suggested control variable, *neighborhood wars*, which captures how many wars were ongoing in contiguous countries in the previous year. The variable has a positive coefficient as expected, but is far from significant ( $p = 0.60$ ) and adds very little to pseudo  $R^2$ .<sup>67</sup> Including this control variable only implies the loss of 34 observations from the analyses and less efficient estimates of the other parameters. I therefore choose to exclude it from the rest of the analyses. Model 2-4 is thus chosen as the *basic model* - my starting point for testing the hypotheses.

## 6.2 Testing the grievance hypotheses

I here test the three hypotheses that have been derived from the grievance explanation. *H1* and *H2* both hold that the poverty-civil war relationship can be explained by the particularly strong grievances that increasing taxation creates in poor countries. These two hypotheses can

---

<sup>65</sup> I also tried adding a square term of  $\log(\text{per capita income})$  to Model 2-3. This was not significant ( $p > 0.10$ ), and did not add much to the fit of the model (pseudo  $R^2 = 0.083$ ). There is thus no significant curvilinearity in the relationship.

<sup>66</sup>  $X=400$  gives a log likelihood of -493.52. For  $X=250$  the log likelihood of -493.53, and for  $X=100$  the log likelihood is -493.57.

notably also be derived from a different theoretical argument; state expansion may cause civil war because it intensifies the power struggle between weak state builders and other domestic social forces. The third hypothesis is derived from a quite different grievance argument; grievances may be particularly strong in poor countries due to the failure of the state to provide public goods that people think themselves entitled to. These hypotheses are tested below.

### **6.2.1 Previous studies on increasing extraction**

The hypotheses on increase of extraction (or state expansion) and civil war are inspired by the arguments of Cohen et al. (1981). These authors perform quantitative empirical tests of their hypothesis that *state expansion in initially weak states leads to increased levels of collective violence*. Their dependent variable, collective violence, is operationalized in two ways; as the average annual number of riots and armed attacks in each country, and as the average annual number of deaths related to these collective violence events. They measure the primary independent variable, *state expansion*, by the annual rate of change in tax ratio. “Strong states” are defined as countries with a tax ratio of at least 16 % and “weak states” as countries with a tax ratio below 16 %. The evidence supports the authors’ hypothesis. A higher rate of state expansion is strongly associated with more collective violence events and related deaths for weak states. For strong states, the relationship is the opposite: a higher rate of state expansion is related to fewer collective violence events, and a somewhat lower number of deaths (Cohen et al. 1981:906-909).

Their results are very fragile, however. Only 31 unspecified countries are considered for the 1950-1965 period. The only plausible confounding variable controlled for is economic growth. Thus, their findings merely show a correlation between the variables for a small and possibly biased sample of countries. Also, this analysis is not very relevant for my purpose, as I focus on the risk of civil war onset, not aggregate levels of collective violence. To my knowledge, the hypotheses on increased extraction and civil war onset have not yet been tested in a quantitative analysis.

---

<sup>67</sup> I also test the variable as a dummy, taking the value 1 if there are one or more wars in contiguous countries, and 0 if no war. This dummy variable has an even higher p-value.

### 6.2.2 The first increase of extraction hypothesis (H1)

The first hypothesis holds (a) that *increase of extraction from a low initial level of extraction leads to a higher risk of civil war onset*, and (b) that *the per capita income effect on risk of civil war onset is considerably weakened by controlling for the interaction between change of extraction and level of extraction*. I do not use Cohen et al.'s (1981) test of *H1a*. They distinguished between “weak states” and “strong states”, and performed a separate analysis for each category. I avoid the difficulty of separating between weak and strong states by testing an interaction term between *tax ratio* and *tax ratio change* (*tax ratio \* tax ratio change*).<sup>68</sup> To reduce collinearity I center the two variables so that their mean value is zero.

The extraction variables have more missing observations than the variables in the basic model (Model 2-4). To assess *H1b*, whether the *per capita income* effect is considerably weakened by controlling for the interaction term between *tax ratio* and *tax ratio change*, we first need to test the basic model without the observations that are missing from the extraction variables. This is done in Model 3-1. The *per capita income* effect is approximately the same as in the more complete sample. The largest change concerns the effect of *oil exporter*, which is not significant even at the 0.10 level. The effects of *ongoing war* and *anocracy* are also reduced, but are still significant at  $\alpha = 0.10$  and  $\alpha = 0.05$  respectively.

In Model 3-2 *tax ratio change* is added to the basic model. Its coefficient has the predicted positive sign, implying that increase of extraction is related to a higher risk of civil war onset. The effect is also significant at the 0.05 level. This can not yet be taken to support *H1* because the hypothesis specifies that the *tax ratio change* effect should be conditional on the level of *tax ratio*. In Model 3-3 I add *tax ratio* to the equation. As predicted by state weakness theory (*H4.4*), *tax ratio* has a negative effect which is significant at  $\alpha = 0.10$ . Thus, a higher level of extraction is related to a lower risk of civil war onset. The effect of *tax ratio change* increases slightly, whereas the *per capita income* coefficient decreases somewhat (by 14 percent).

---

<sup>68</sup> The two component variables are described in section 5.2.1.

**Table 3.** Test of H1 and H2: Estimating the conditional effect of change in tax ratio on the risk of civil war onset.

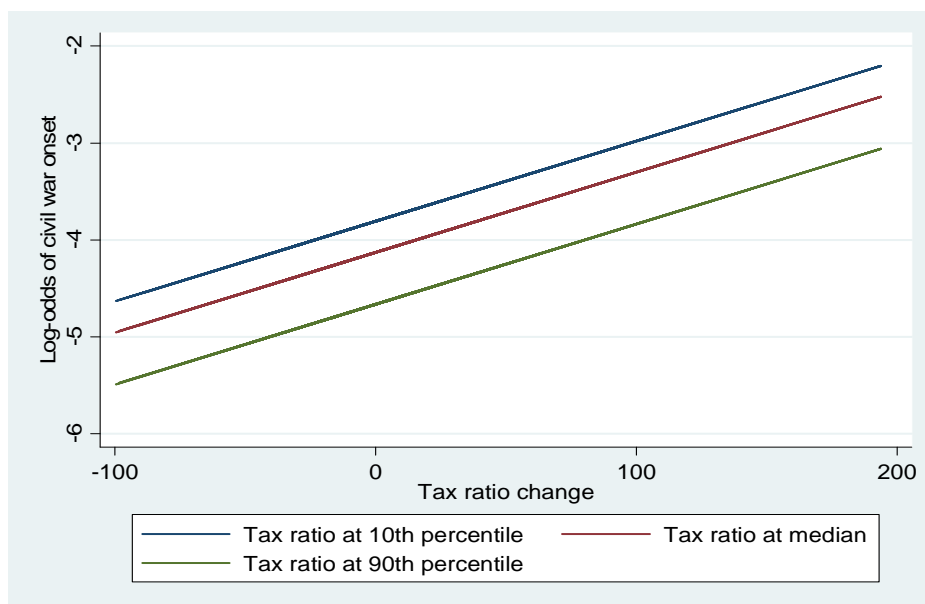
Variable	Model 3-1	Model 3-2	Model 3-3	Model 3-4	Model 3-5	Model 3-6
Tax ratio change <sub>d, e</sub>		0.0077** (0.0039)	0.0083** (0.0037)	0.012*** (0.0045)	0.011** (0.0047)	0.0094** (0.0040)
Tax ratio <sub>a, d</sub>			-3.57* (2.11)	-4.00* (2.15)		-4.90** (2.33)
Tax ratio change <sub>d</sub> * Tax ratio <sub>d</sub>				0.050 (0.043)		
Tax ratio change <sub>c, d</sub> * Income per capita <sub>a, b, d</sub>					0.0051 (0.0044)	
Change in military personnel per capita						0.0027 (0.0033)
Economic growth <sub>a</sub>						-0.026 (0.022)
Income per capita <sub>a, b, d</sub>	-0.72*** (0.15)	-0.71*** (0.16)	-0.60*** (0.17)	-0.58*** (0.17)	-0.72*** (0.16)	-0.53*** (0.18)
Population <sub>b</sub>	0.28*** (0.089)	0.28*** (0.089)	0.27*** (0.089)	0.27*** (0.089)	0.29*** (0.089)	0.26*** (0.095)
Ongoing war <sub>a</sub>	-0.67* (0.35)	-0.68* (0.35)	-0.73** (0.36)	-0.71** (0.36)	-0.67* (0.35)	-0.74** (0.37)
Oil exporter <sub>c</sub>	0.35 (0.38)	0.36 (0.38)	0.34 (0.39)	0.31 (0.39)	0.33 (0.38)	0.16 (0.41)
Anocracy <sub>a, c</sub>	0.68** (0.28)	0.73*** (0.28)	0.68** (0.28)	0.68** (0.28)	0.72*** (0.28)	0.69** (0.29)
Brevity of peace	0.49 (0.30)	0.50* (0.30)	0.46 (0.30)	0.45 (0.30)	0.49* (0.30)	0.57* (0.30)
Constant	-6.64*** (0.85)	-6.71*** (0.85)	-6.59*** (0.85)	-6.76*** (0.86)	-7.29*** (0.86)	-6.64*** (0.92)
Pseudo R <sup>2</sup>	0.086	0.092	0.097	0.099	0.094	0.11
Log likelihood	-281.33	-279.57	-278.04	-277.44	-278.97	-260.19
Observations	3767	3767	3767	3767	3767	3510
Onset Events	60	60	60	60	60	57

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous, d = Values are centered, e = Outliers above the value 200 are excluded.

A possible interaction effect can be assessed by plotting the estimated risk of civil war onset (in log-odds) as a function of *tax ratio change* for different levels of *tax ratio*. The resulting graphs are shown in Figure 10. The blue line represents the estimated relationship between *tax ratio change* and the risk of civil war onset when *tax ratio* is at its 10<sup>th</sup> percentile, the red line when *tax ratio* is at its median, and the green line when *tax ratio* is at its 90<sup>th</sup> percentile. We see that the three lines are approximately parallel, indicating that there is no considerable interaction effect. Contrary to the hypothesis' prediction, the slope of *tax ratio change* is not markedly steeper at a lower level of *tax ratio*.

**Figure 10.** Estimated risk of civil war onset (in log-odds) by tax ratio change and tax ratio, based on Model 3-3.



In Model 3-4 I add the interaction term between *tax ratio* and *tax ratio change*. Its coefficient shows how much the effect of one component variable changes when the other component increases with one unit. Contrary to the prediction, the interaction term's coefficient is positive. The estimate is small, as suggested by the graphs in Figure 10. It is not significant at the 0.10 level. I may conclude that the effect of *tax ratio change* is not stronger at lower levels of *tax ratio*.<sup>69</sup>

It is not likely that imprecision due to multicollinearity affects these results appreciably. As explained in Section 4.2, the level of multicollinearity can be estimated by regressing each independent variable on the other independent variables. On the basis of this a variance inflation factor (VIF) can be estimated, which indicates the extent to which multicollinearity creates greater variance in the estimates. A common rule of thumb is that VIF-values larger than 10.0 are problematic (Kleinbaum et al. 1988:210). However, this also depends on how many observations are included. With few observations, VIF-values below 10 can also contribute to making estimates insignificant (see Section 4.2). In this case, the VIF-values are quite small for all the variables ( $VIF < 1.5$ ), indicating a very low degree of multicollinearity (Appendix 2.1).

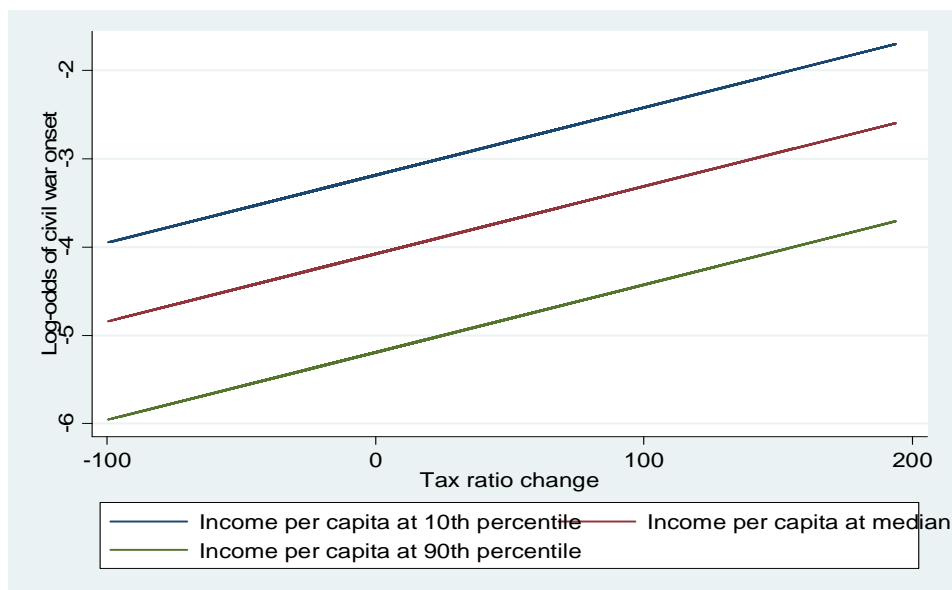
<sup>69</sup> I also tested for curvilinearity by adding a square term of *tax ratio change* to Model 3-4. This was not significant ( $p=0.132$ ).

The second part of *H1* holds that the *per capita income* effect should be weakened by controlling for the interaction between *tax ratio change* and *tax ratio*. This is not supported by the evidence either. The reduction of the *per capita income* effect in Model 3-3 was due to the control for *tax ratio*. Controlling for the interaction between *tax ratio change* and *tax ratio* did not appreciably affect the *per capita income* coefficient. Thus, despite the interesting finding that *tax ratio change* is positively related to the risk of civil war onset, *H1* can be dismissed.

### 6.2.3 The second increase of extraction hypothesis (H2)

The second hypothesis holds that *increase of extraction leads to a higher risk of civil war onset*, and the effect is stronger the poorer is the country. Hence, it predicts that *tax ratio change* should have a stronger effect at low levels of *per capita income*. This can be investigated by plotting the estimated risk (in log-odds) of civil war onset from Model 3-2 as a function of *tax ratio change* for different levels of *per capita income*. The plot is shown in Figure 11. The blue line represents the estimated relationship between *tax ratio change* and the log-odds of civil war onset when *per capita income* is at its 10<sup>th</sup> percentile, the red line when *per capita income* is at its median, and the green line when *per capita income* is at its 90<sup>th</sup> percentile. Like in the previous figure, the three lines are approximately parallel, indicating no discernible interaction effect.

**Figure 11.** Estimated risk of civil war onset (in log-odds) by tax ratio change and income per capita, based on Model 3-2.



In Model 3-5 I add an interaction term between *per capita income* and *tax ratio change* (*per capita income* \* *tax ratio change*) to Model 3-2. Contrary to the prediction, the interaction term coefficient has a positive sign. The estimate is small, as Figure 11 indicated, and it is not significant at the 0.10 level. There are no signs that the high p-value is due to multicollinearity. A large proportion of the variables' variance is independent of the other variables (Appendix 2.2). *H2* thus receives no empirical support. I may conclude that the poverty-civil war relationship appears not be explained by increased extraction (or state expansion) in poor countries with weak states.

#### 6.2.4 Assessing the effect of increasing extraction

Although the two hypotheses were not supported, the finding of a positive *tax ratio change* effect is interesting per se. I therefore investigate somewhat further whether there may actually be a causal relationship. I test two alternative hypotheses. First, *economic growth* may be a confounding variable here, inducing the positive correlation found between *tax ratio change* and civil war onset. It is possible that the *tax ratio* often increases during economic downturns. This may happen if governments do not adjust tax rates according to economic fluctuations. Thus, economic crisis may be the real cause of civil war, whereas the effect of *tax ratio change* is spurious. I test this in Model 3-6 by controlling for *economic growth*. The hypothesis is not supported. *Economic growth* does not reach significance, and the positive effect of *tax ratio change* is still significant at the 0.05 level.

I also try to assess whether the *tax ratio change* effect may be due to endogeneity. A plausible endogeneity hypothesis is that a brewing armed rebellion causes the government to increase extraction efforts in order to invest more in counterinsurgency. If this were true, we should expect that the significant effect of *tax ratio change* should disappear when controlling for change in counterinsurgent capacity.<sup>70</sup> Counterinsurgent capacity is proxied by change in military personnel per 1000 inhabitants ("*change in military personnel per capita*") (see Section 5.4.1 and 6.3.1).<sup>71</sup> This variable is included in Model 3-6. The results show that the effect of *tax ratio change* remains positive and significant, whereas *change in military*

<sup>70</sup> I have not seen this method of reducing the endogeneity problem applied before. This should be a useful method when an instrumental variable approach is out of reach.

<sup>71</sup> It is measured by the percent change in military personnel per capita from 3 years before to 1 year before the observation, just like *tax ratio change*. I drop outliers that display more than 300 percent change (24 observations), as these implausible values make the standard deviation extremely high.



*personnel per capita* is far from significant ( $p = 0.41$ ). Hence, my tests suggest that increase of extraction may be causally related to civil war onset. This should be further investigated in future research.

The reliability of the results in this section can be questioned. Clearly, data on taxation are politically sensitive (Jackman 1993:52). Governments may find it in their interest to report too high or too low figures. Only governments facing weak institutional checks and fiscal review have the opportunity to fix the figures, however. These factors may be correlated with economic development. It is plausible that rich countries on average have more checks on government than poor countries. This implies that measurement error in the data on *tax ratio* may have a systematic component. The direction of bias from the systematic error component is uncertain for the tax ratio. Governments or other state agencies may find that much tax revenue shows strength, and want to exaggerate numbers. On the other hand, corrupt officials may underreport tax revenues in order to spend them privately.

The systematic measurement error is probably smaller when looking at *tax ratio change*. The degree to which governments report wrong figures probably does not change systematically over consecutive years within one country. On the other hand, the problem of *random measurement error* is probably greater for this variable. I have already noted that some observations have implausibly high *tax ratio change* values. Outliers having more than 200 percent tax ratio increase were excluded from my analyses. There is a certain possibility that coincidence and random error causes the positive effect. This should be investigated thoroughly. I have only managed to do some scattered checks. These indicated that the results are not merely coincidental.<sup>72</sup> The random measurement error is more likely to bias the coefficient estimate toward no relationship, as explained by King et al. (1994:163-168). Still, the significant reliability problem suggests caution when making inferences.

---

<sup>72</sup> None of the extreme outlier country-year observations have civil war onsets. The highest value of *tax ratio change* for an onset observation was 93.3 percent (Senegal 1989). In most cases of onset I looked at, the tax ratio had increased systematically several years before the onset. For example, the civil war onset in Argentina in 1973 corresponded to a 34 percent tax ratio increase. The "Bureaucratic-Authoritarian State" had then been increasing extraction greatly over several years - from a tax ratio of 0.05 in 1965 to 0.14 in 1973. Similarly, in Indonesia there was a clear state expansion trend from the early 1960s to 1980, in which period there were two civil war onsets.

### 6.2.5 The primary schooling hypothesis (H3)

The third grievance hypothesis holds (a) that *lower levels of public goods provision, such as public education, are related to a higher risk of civil war onset*, and (b) that *the per capita income effect on risk of civil war onset is considerably weakened by controlling for public goods provision*. Gross total primary school enrollment (*primary schooling*) was chosen as a proxy for basic public goods provision in Section 5.3.2.

Thyne (2006) has made the most comprehensive quantitative study of education and civil war. He tests the relationship between several education variables, including *primary schooling*, and the risk of civil war onset, using the Fearon-Laitin civil war measure. *Primary schooling* has a negative and significant effect in his models, even when controlling for *per capita income*. The *per capita income* effect is also quite considerably weakened when controlling for *primary schooling*, as H3 predicts (Thyne 2006:741-742).

Hypothesis 3b concerns how much the *per capita income* effect is reduced by controlling for *primary schooling*. I investigate this by comparing a model including the education variable and a model excluding it. This only gives meaning if the two models look at the same observations. Model 4-1 shows the results of the basic model when excluding all the observations that are missing from the primary schooling variable. In Model 4-2 I include *primary schooling*. Its coefficient has the predicted negative effect, and is significant at the 0.10 level. Thus, a lower level of primary school enrollment is related to a higher risk of civil war, even when *per capita income* is held constant. The *per capita income* coefficient is quite considerably reduced (by 31 percent), and is now significant only at the 0.05 level.

In Model 4-3 *per capita income* is excluded. *Primary schooling* then becomes significant at the 0.01 level. The lower p-value is partly due to a reduction of its standard error. This indicates that the correlation between *per capita income* and *primary schooling* made the estimates of their individual effects less precise.<sup>73</sup> However, the reduced p-value of *primary schooling* in Model 4-3 owes more to the increase of its coefficient, which is almost doubled from Model 4-2. This indicates that *per capita income* is a causally more important variable, since much of the effect of *primary schooling* disappears when holding *per capita income* constant. A comparison of the log likelihood of Model 4-1 and Model 4-3 points in the same

---

<sup>73</sup> Their coefficients in Model 4-2 are correlated by  $r = -0.61$ .

direction. The model excluding *primary schooling* has a somewhat lower log likelihood (-263.40) than the model excluding *per capita income* (-264.57).

**Table 4.** Test of H3: Estimating the effect of primary schooling on the risk of civil war onset.

Variable	Model 4-1	Model 4-2	Model 4-3
Primary schooling <sub>a</sub>		-0.010* (0.0059)	-0.019*** (0.0047)
Income per capita <sub>a, b</sub>	-0.64*** (0.16)	-0.44** (0.19)	
Population <sub>b</sub>	0.31*** (0.090)	0.34*** (0.095)	0.35*** (0.095)
Ongoing war <sub>a</sub>	-0.71* (0.36)	-0.70* (0.36)	-0.59 (0.36)
Oil exporter <sub>c</sub>	0.13 (0.41)	0.16 (0.41)	0.012 (0.40)
Anocracy <sub>a, c</sub>	1.092*** (0.28)	1.093*** (0.28)	1.15*** (0.28)
Brevity of peace	0.63** (0.30)	0.72** (0.31)	0.89*** (0.30)
Constant	-7.00*** (0.87)	-6.55*** (0.91)	-6.24*** (0.91)
Pseudo R <sup>2</sup>	0.094	0.099	0.090
Log likelihood	-263.40	-261.85	-264.57
Observations	3463	3463	3463
Onset Events	57	57	57

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous.

Still, the results at least partly support the hypothesis. *Primary schooling* has a negative effect on the risk of civil war onset which is significant at the 0.10 level. Moreover, it accounts for a considerable part of the *per capita income* effect, as predicted by H3b. The argument that a fraction of the causal *per capita income* effect is indirect, going through public goods provision, can thus not be dismissed. These results may be cited in support of the grievance explanation. However, as I discussed in chapter three, this grievance argument is theoretically disputed. It rests on the quite fragile assumption that poor people generally expect the state to provide a certain level of primary education. We must therefore be careful about making causal inferences. Concluding very cautiously, the grievance explanation can at least not be rejected on the basis of the findings.

I have argued that the school enrollment data are more reliable than data on education expenditure. Still, there might be some systematic measurement error. Governments may want to report too high figures for political reasons. Plausibly, this happens more often in poor countries with low enrollment rates. Officials in poor countries may have a greater interest in boosting the numbers (for example to show progress to aid donors). If this holds, it

would mean that the variance in the independent variable declines. This implies a bias of the statistical results toward no relationship with the dependent variable. It could also mean that the measure shows less than true correlation with *per capita income*. I do not know the extent of the reliability problem, but if there is measurement error it most likely biases the estimates toward no causal relationship. Low reliability should therefore not be a problem for my conclusion that the hypothesis is supported.

### **6.2.6 Summary of results for the grievance hypotheses**

The two first grievance hypotheses concerning increase of extraction were not supported by the evidence. I found that tax ratio increase was positively related to the risk of civil war onset. However, the effect of tax ratio increase was not stronger in countries with an initially lower level of extraction, and neither in poorer countries. Thus, it could not account for the poverty-civil war relationship, as the two first grievance hypotheses predicted. Notably, these two hypotheses could also be derived from a different theoretical framework which emphasizes conflicts of interest between state builders and rival domestic social organizations. This argument is similarly weakened by the results.

The third grievance hypothesis concerning lack of public goods provision in poor countries received some empirical support. The argument that the *per capita income* effect is partly indirect, going through public goods provision, could at least not be disconfirmed. I have been cautious about making strong inferences, as the theoretical argument underlying the third grievance hypothesis rests on a highly uncertain assumption.

## **6.3 Testing the state weakness hypotheses**

In chapter three I derived five hypotheses from the state weakness explanation. Four of these concerned variables thought to capture state capacity in different ways (*H4.1*, *H4.2*, *H4.3*, and *H4.4*). The fifth hypothesis concerns *change* in state capacity, or particularly, in government revenues. When testing the tax ratio hypothesis (*H4.4*), I also assess the converse greed hypothesis concerning tax ratio (*H8.5*). All these hypotheses are tested below.

### 6.3.1 The military personnel hypothesis (H4.1)

The first hypothesis specifically related to the state weakness explanation concerns military strength. Hypothesis 4.1 holds (a) that *the number of military personnel per capita is negatively related to the risk of civil war onset*, and (b), that *the per capita income effect on risk of civil war onset is considerably weakened by controlling for military personnel per capita*. The explanatory variable is measured by military personnel per 1000 inhabitants. I refer to it as *military personnel per capita* for the sake of simplicity.

Hegre and Sambanis (2006:21) find that *military personnel as a share of the population* is negatively and robustly related to civil war onset. In the same study, they find that *military expenditure as a percentage of GNP* has a *positive* effect on the risk of civil war. The latter is also found by Henderson and Singer (2000:282). I have argued, however, that military expenditure is a very bad proxy for state counterinsurgent capacity. The validity of *military personnel per capita* can also be questioned, but it should clearly be more reliable than the expenditure data (Herbst 2004:364).

In Model 5-1 I estimate the basic model again, excluding the observations missing from the military personnel variable. I include *military personnel per capita* in Model 5-2. Its coefficient has the predicted negative effect, and is significant at the 0.10 level. Hence, a higher number of military personnel per capita is related to a lower risk of civil war.<sup>74</sup> The second part of the hypothesis is not supported, however. The *per capita income* effect is only reduced by 9 percent when controlling for *military personnel per capita*. This is below my threshold for what is “considerable” (20 %). Differences in military strength, as measured here, appear not to explain a significant part of the poverty-civil war relationship.

---

<sup>74</sup> There is not much collinearity involved, making it unnecessary to estimate a model without *per capita income*. 96 percent of *military per capita*'s variation is independent (1/VIF), and the *per capita income* and *military personnel per capita* coefficients are correlated by  $r = -0.33$ .

**Table 5.** Test of H4.1 and H4.2: Estimating the effect of primary schooling on the risk of civil war onset.

Variable	Model 5-1	Model 5-2	Model 5-3	Model 5-4	Model 5-5
Military personnel per capita <sub>a, b</sub>		-0.19* (0.11)			
Road density				-0.92 (0.70)	-1.91** (0.76)
Population density <sub>b</sub>				0.18 (0.13)	0.26** (0.12)
Income per capita <sub>a, b</sub>	-0.68*** (0.11)	-0.62*** (0.12)	-0.58*** (0.15)	-0.44** (0.17)	
Population <sub>b</sub>	0.25*** (0.070)	0.25*** (0.071)	0.34*** (0.089)	0.29*** (0.096)	0.27*** (0.093)
Ongoing war <sub>a</sub>	-0.82*** (0.30)	-0.81*** (0.30)	-0.69* (0.37)	-0.68* (0.37)	-0.53 (0.36)
Oil exporter <sub>c</sub>	0.75*** (0.27)	0.76*** (0.27)	0.42 (0.36)	0.45 (0.37)	0.25 (0.36)
Anocracy <sub>a, c</sub>	0.78*** (0.20)	0.75*** (0.20)	0.80*** (0.26)	0.75*** (0.26)	0.73*** (0.26)
Brevity of peace	0.32 (0.24)	0.37 (0.24)	0.45 (0.29)	0.34 (0.30)	0.39 (0.30)
Constant	-6.43*** (0.66)	-6.18*** (0.68)	-7.36*** (0.86)	-6.30*** (0.70)	-5.88*** (1.09)
Pseudo R <sup>2</sup>	0.080	0.083	0.074	0.078	0.069
Log likelihood	-488.26	-486.63	-309.55	-308.06	-311.23
Observations	6116	6116	4392	4392	4392
Onset Events	108	108	64	64	64

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous.

### 6.3.2 The road density hypothesis (H4.2)

I have argued that road density can be used as a proxy for the state's reach into rural areas and for the conduciveness of the physical environment to insurgency. Thus, it can be used to evaluate both the state capacity and the ecological explanation. This is not surprising, as the two explanations build on a quite similar theoretical framework (see Section 3.3). The hypothesis (H4.2) holds (a) that *a lower road density is associated with a higher risk of civil war onset*, and (b) that *the per capita income effect on risk of civil war onset is considerably weakened by controlling for road density*. I measure *road density* by the total length of roads per 1000 square kilometers.

This measure has to my knowledge not been tested in any quantitative, cross-national analysis of civil war. Buhaug and Rød (2006) include *road density* in a disaggregated analysis of *local determinants* of African civil wars. They find that those areas where *conflicts over territory* break out have a significantly lower *road density* than the average for the country. On the other hand, the areas where *conflicts over state power* begin have a somewhat higher *road density* than the average for the country. These findings are interesting, but not particularly

relevant for my study, which does not separate between civil wars over territory and over government, and which concentrates on the national level.

Model 5-3 shows the results of a logistic regression of civil war onset on the basic model when excluding the observations missing from the road variable. In Model 5-4 I add *road density* to the basic model. An additional control variable is also included: *population density*. In more densely populated areas there will be a greater economic demand for roads. As expected, *road density* increases with *population density* ( $r = 0.54$ ). I am mainly interested in variations of infrastructure which are due to state capacity and public policy, not due to economic forces of demand and supply. Thus, I want to look at *road density* keeping *population density* at a constant level. *Population density* is logged to reduce its skewness.<sup>75</sup>

The *road density* coefficient has the predicted negative sign, but the effect is not significant even at the 0.10 level. Interestingly, the *per capita income* effect is weakened by 24 percent. The relatively high p-value of *road density* may be partly due to collinearity making the estimates of individual coefficients imprecise and standard errors high. The VIF-value for *road density* is 1.97, showing some, but not high multicollinearity (Appendix 2.3). In Model 5-5 I exclude *per capita income* from the analysis. The negative effect of *road density* then turns up as significant at the 0.05 level. This is entirely due to the increase of its coefficient, which is more than doubled. Its standard error actually increases when excluding *per capita income*. This indicates that the non-significance of *road density* in Model 5-4 is not due to collinearity causing low precision of estimates. It is primarily due to *road density* having no considerable effect on the risk of civil war when keeping *per capita income* at a constant level.

The results are thus mixed when it comes to assessing *H4.2*. Higher *road density* is significantly related to a lower risk of civil war onset if we do not control for *per capita income*. Controlling for *per capita income* the effect is still negative, but not significant. As *H4.2b* suggests, when controlling for *road density*, the *per capita income* effect is considerably reduced. The relative lack of roads in poor countries may thus be a (minor) part of the explanation for the poverty-civil war relationship.

---

<sup>75</sup> I decide not to log-transform *road density* because it is not as skewed. Log-transforming *road density* also reduces the fit of the model considerably.

### 6.3.3 The government share of consumption hypothesis (H4.3)

The third proxy I have suggested for state strength is the government share of consumption (“*government share*”). I argued that this can not be interpreted as an encompassing measure of state strength, but that it should capture some of the state’s economic muscle (see Section 3.2.8). The state weakness explanation may thus suggest *H4.3*, which holds (a) that a higher government share of consumption is associated with a lower risk of civil war onset and (b) that the per capita income effect on risk of civil war onset is considerably weakened by controlling for government share of consumption.

Ross (2006:288) performs the only test I have seen of *government share* in published quantitative conflict studies. He finds that it has no significant effect on the risk of civil war onset using the Fearon-Laitin civil war data and the various Uppsala/PRIO armed conflict measures. Using a civil war measure from Sambanis (2004b) he finds a *positive* effect which is significant at  $\alpha = 0.10$ . This is the opposite of what *H4.3* predicts. Unfortunately, he does not report where he gets his data for government share of consumption from. As I have discussed, these measures differ widely, and the WDI 2003 measure is probably the most appropriate for my purpose (Section 5.4.3). Moreover, Ross does not report important details about the analysis such as the number of observations and onsets included.

I begin with a logistic regression of the basic model excluding observations missing from *government share*. The results are shown in Model 6-1. *Government share* is added to the equation in Model 6-2. Its coefficient is negative as expected, but far from significant even at the 0.10 level. The non-significance of the effect can not be explained by imprecise estimation due to multicollinearity. The VIF-value of *government share* is 1.15, indicating no discernible degree of multicollinearity (Appendix 2.4). Hence, *H4.3a* is not supported. A higher government share of consumption is not significantly related to a lower risk of civil war. Moreover, the *per capita income* effect is barely weakened by controlling for *government share*. This suggests that variations in government share of consumption can not explain the poverty-civil war relationship. Hence, the second part of the hypothesis should also be dismissed.



**Table 6.** Test of H4.3 and H5: Estimating the effect of government share and decrease of government share on the risk of civil war onset.

Variable	Model 6-1	Model 6-2	Model 6-3	Model 6-4	Model 6-5
Government share <sub>a</sub>		-0.019 (0.027)			
Decreasing revenues <sub>c</sub>				0.17 (0.27)	0.16 (0.32)
Tax ratio <sub>a, d</sub>					-2.65 (3.08)
Decreasing revenues <sub>c</sub> * Tax ratio <sub>a, d</sub>					0.12 (4.04)
Income per capita <sub>a, b</sub>	-0.74*** (0.15)	-0.72*** (0.16)	-0.76*** (0.16)	-0.76*** (0.16)	-0.68*** (0.18)
Population <sub>b</sub>	0.31*** (0.087)	0.29*** (0.089)	0.33*** (0.091)	0.33*** (0.091)	0.30*** (0.096)
Ongoing war <sub>a</sub>	-0.90** (0.36)	-0.90** (0.36)	-0.83** (0.37)	-0.82** (0.37)	-0.71* (0.37)
Oil exporter <sub>c</sub>	0.55 (0.35)	0.59* (0.35)	0.33 (0.39)	0.33 (0.39)	0.13 (0.43)
Anocracy <sub>a, c</sub>	1.04*** (0.26)	1.02*** (0.26)	0.88*** (0.28)	0.88*** (0.28)	0.75** (0.30)
Brevity of peace	0.56* (0.29)	0.54* (0.30)	0.66** (0.31)	0.66** (0.31)	0.54* (0.32)
Constant	-6.98*** (0.82)	-6.60*** (0.99)	-7.15*** (0.87)	-7.23*** (0.88)	-6.98*** (0.93)
Pseudo R <sup>2</sup>	0.11	0.11	0.11	0.11	0.10
Log likelihood	-295.10	-294.86	-259.18	-258.99	-238.41
Observations	3956	3956	3522	3522	3221
Onset Events	65	65	57	57	52

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous, d = Values are centered.

I have noted that the *government share of consumption* data probably contain large errors of measurement. Most of this seems to be random, due to varying standards of reporting in different countries and little control of reliability. This will increase the standard error and most likely bias the estimates of its effect on the risk of civil war in the direction of no relationship (King et al. 1994:163-168). There may also be a systematic error component here. As argued before, poor countries with weak states are more likely to lack checks and fiscal reviews than wealthy states. Governments probably find it in their interest to report spending on various public purposes to be *higher* than it really is. If this line of thought is correct, it biases the estimate of a causal effect toward zero. Thus, we cannot with high degree of certainty conclude that the hypothesis is wrong on the basis of the above results.

### 6.3.4 The decreasing revenues hypothesis (H5)

In chapter three I discussed the argument that many civil wars in poor countries can be explained by a decrease of government revenues in weak states (Section 3.2.10). *H5* holds (a) that *decreasing government revenues in weak states are related to a higher risk of civil war onset*, and (b) that *the per capita income effect on the risk of civil war onset is*

*considerably weakened by controlling for the interaction between state capacity and change in revenues.* Change in government revenues was operationalized as change in the government share of consumption (“*government share change*”) in Section 5.4.3. I have not seen any test of this variable before in the quantitative literature.

The hypothesis does not predict that any increase in *government share* is related to a lower risk of civil war. It only specifies that clearly negative values of *government share change* in weak states should be related to a higher risk of civil war. Thus, the effect of *government share change* is not predicted to be constant at all levels of *government share change*. Such a non-linear relationship between the independent and the dependent variable can well be captured by a dummy variable. I create a *decreasing revenues* dummy taking the value 1 if *government share* decreases from three to one year before the observation, and 0 if it is constant or increasing. This variable is included in Model 6-4. Its coefficient is positive, as predicted, but far from significant at the 0.10 level. Hence, decreasing revenues is not significantly related to a higher risk of civil war in general.<sup>76</sup>

*H5* emphasizes that the effect is conditional on state strength, however. Decreasing revenues should be positively related to civil war onset particularly in weak states. Finding a single, good proxy for state weakness is probably not possible. I still need to choose one here to assess the decreasing revenues hypothesis. The *tax ratio* may be the best available candidate. In Model 6-5 I add an interaction term between *tax ratio* and *decreasing revenues* to test the hypothesis. The hypothesis is not supported. Decreasing revenues is still far from significant at the 0.10 level. The interaction term has a positive sign, contrary to the prediction. This indicates that one unit increase of tax ratio *increases* the positive effect of decreasing revenues. Its coefficient is far from significant at the 0.10 level, however. Thus, the results do not support the hypothesis. Decreasing revenues in countries with a weak state is not significantly related to a higher risk of civil war. Consequently, it can not explain the poverty-civil war relationship either.<sup>77</sup>

It must be noted that the problem of random errors in the *government share* data is magnified when looking at *government share change*. The figures reported tend to change markedly

---

<sup>76</sup> I tested an alternative specification of the decreasing revenues dummy, taking the value 1 if government share decreased by more than 10 percent. The results were even less supportive of the hypothesis.

from year to year, creating a high standard deviation. Also, there are many onset observations missing from this variable, which decreases the possibility for making inferences. Thus, the hypothesis should perhaps not be too readily dismissed on the basis of these findings.

### 6.3.5 The tax ratio hypotheses (H4.4 and H8.5)

In chapter three I argued that the state capacity and the greed theory may give opposite predictions concerning the relationship between state extraction and the risk of civil war onset. The state capacity theory holds that a higher level of extraction should be related to a lower risk of civil war because it indicates state strength. The greed theory, on the other hand, suggests that a high level of extraction may be related to a higher risk of civil war because it implies lower economic opportunities in the regular labor market. The latter argument may not be correct in its unconditional form, as “greed scholars” also acknowledge that higher state capacity gives less opportunity for rebellion (e.g. Collier and Hoeffler 2004).

I thus derived *H8.5* from the greed theory, which holds that extraction level (measured by the tax ratio) is positively related to the risk of civil war onset *when controlling for state counterinsurgent capacity*. From state capacity theory I suggested *H4.4*, which predicts (a) that *a higher tax ratio is associated with a lower risk of civil war onset*, and (b) that *the per capita income effect on risk of civil war onset is considerably weakened by controlling for tax ratio*. I have not seen these hypotheses tested in any quantitative study.

I first test the state capacity hypothesis. Model 7-1 shows the logistic regression of onset on the basic model when excluding the observations missing from *tax ratio*. In Model 7-2 *tax ratio* is added to the basic model. Its coefficient is negative as predicted by *H4.4*, implying that a higher *tax ratio* is associated with a lower risk of civil war onset. The effect is not significant even at the 0.10 level, however. There is no sign of high multicollinearity causing this insignificance (Appendix 2.1). The second part of the hypothesis is not supported either. The effect of *per capita income* is not considerably reduced by controlling for *tax ratio*. I may thus conclude that *H4.4* is not supported. The effect of *per capita income* on the risk of civil war appears not to be explained by a lower *tax ratio* in poor countries.

---

<sup>77</sup> The per capita income coefficient is not weakened at all when controlling for decreasing revenues in Model 6-4. Model 6-5 should not be compared to Model 6-3 as they include different observations.

**Table 7. Test of H4.4 and H8.5: Estimating the effect of tax ratio on the risk of civil war onset.**

Variable	Model 7-1	Model 7-2	Model 7-3
Tax ratio <sub>a</sub>		-1.97 (1.93)	-2.02 (1.95)
Military personnel per capita <sub>a, b</sub>			-0.057 (0.17)
Income per capita <sub>a, b</sub>	-0.77*** (0.15)	-0.70*** (0.17)	-0.68*** (0.18)
Population <sub>b</sub>	0.28*** (0.087)	0.27*** (0.087)	0.25*** (0.089)
Ongoing war <sub>a</sub>	-0.73** (0.35)	-0.74** (0.35)	-0.72** (0.35)
Oil exporter <sub>c</sub>	0.56 (0.35)	0.55 (0.35)	0.56 (0.35)
Anocracy <sub>a, c</sub>	0.60** (0.27)	0.57** (0.27)	0.56** (0.27)
Brevity of peace	0.37 (0.29)	0.35 (0.30)	0.37 (0.30)
Constant	-6.64*** (0.83)	-6.28*** (0.89)	-6.00*** (0.92)
Pseudo R <sup>2</sup>	0.085	0.086	0.086
Log likelihood	-302.63	-302.08	-300.84
Observations	4154	4154	4051
Onset Events	64	64	64

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous

I next consider the greed hypothesis, *H8.5*. Like above, I use *military personnel per capita* as a proxy for counterinsurgent capacity. This is included as a control variable in Model 7-3. The coefficient of *tax ratio* barely changes and is still clearly negative, though not significant. Thus, although variation in extraction cannot explain the poverty-civil war relationship, the evidence is more in favor of state capacity theory than greed theory. The results clearly dismiss the greed hypothesis that a higher level of extraction increases the risk of civil war.

### 6.3.6 Summary of results for the state weakness hypotheses

Most of the state weakness hypotheses received little or no empirical support. In accordance with *H4.1a*, I found that *military personnel per capita* was negatively related to the risk of civil war onset. However, its effect was largely independent of the poverty-civil war relationship. *H4.1b* could thus be dismissed. The results were more favorable to *H4.2*. The *per capita income* effect was considerably reduced when controlling for *road density*. However, the negative effect of *road density* was not significant when controlling for *per capita income*. Thus, the results partly supported the hypothesis. *H4.3* found absolutely no empirical support, as the effect of *government share of consumption* was far from significant, and did not weaken the *per capita income* effect. The same was the case for *H4.4*, concerning the *tax ratio*, and for *H5*, which

predicted that decreasing revenues in weak states could help account for the poverty-civil war relationship.

The results thus lend very little support to the state weakness explanation. This must be qualified by the dubious validity and reliability of the state capacity proxies. I have tried to compensate for this problem by testing several different proxies more specifically related to state capacity than is *per capita income*. If state weakness could really account for most of the poverty-civil war relationship, we should expect at least some of the hypotheses to be supported. These results thus do not speak in favor of the state weakness explanation. This should not be taken as evidence that *state weakness theory* in its entirety is wrong, or that state weakness does not increase the risk of civil war. The finding of a negative effect of *military personnel per capita* may suggest that state weakness *is* important, but not as an explanation for the poverty-civil war relationship.

## **6.4 Testing the ecological hypotheses**

In Section 3.3 I derived three hypotheses from the ecological explanation. The risk of civil war should decrease with the proportion of people living in urban areas (*H6*), with the extensiveness of telecommunication infrastructure (*H7*), and with increasing road density (*H4.2*). These three factors may all indicate how conducive the physical and social environment is to insurgency and guerilla warfare. I test only the two first hypotheses here, as the road density hypothesis was tested above in Section 6.3.2.

### **6.4.1 The urban settlement hypothesis (H6)**

The first ecological explanation holds that poor countries are more prone to civil war because they tend to have a rural settlement structure, which is favorable to insurgency. The corresponding hypothesis (*H6*) predicts (a) that *a higher proportion of people living in urban areas is related to a lower risk of civil war onset*, and (b) that *the per capita income effect on risk of civil war onset is considerably weakened by controlling for the proportion of people living in urban areas*.

Kocher (2004) finds that the percentage of inhabitants living in urban areas (“% urban”) is strongly and negatively related to the risk of civil war onset.<sup>78</sup> The effect of the variable is significant at  $\alpha = 0.01$  in his model, even when controlling for *per capita income*. Notably, he finds that the *per capita income* effect reverses and becomes positive when controlling for % urban. Thus, according to his results, a higher *per capita income* is related to a higher risk of civil war for countries having an equally urban settlement structure (Kocher 2004:68-69).

I begin with a logistic regression of onset on the basic model when excluding the observations missing from % urban. The results are shown in Model 8-1. In Model 8-2 I have included % urban. The variable has a negative coefficient as predicted, and is significant at the 0.10 level. The *per capita income* coefficient is almost halved, clearly supporting *H6b*. In contrast to Kocher (2004), however, I find that the *per capita income* effect has not changed its sign. It is still negative and almost significant at the 0.10 level ( $p^*=0.10$ ). Settlement pattern thus appears not to explain the entire *per capita income* effect, as he claims.

**Table 8.** Test of *H6* and *H7*: Estimating the effect of settlement structure and telecommunication infrastructure on the risk of civil war onset.

Variable	Model 8-1	Model 8-2	Model 8-3	Model 8-4	Model 8-5	Model 8-6
% Urban		-0.018* (0.010)	-0.032*** (0.0059)			
Telephones <sub>a, b</sub>					-0.18 (0.15)	-0.34*** (0.072)
Income per capita <sub>a, b</sub>	-0.72*** (0.13)	-0.38 (0.23)		-0.69*** (0.15)	-0.38 (0.30)	
Population <sub>b</sub>	0.25*** (0.077)	0.25*** (0.076)	0.25*** (0.075)	0.32*** (0.081)	0.32*** (0.080)	0.33*** (0.079)
Ongoing war <sub>a</sub>	-0.79** (0.32)	-0.79** (0.32)	-0.75** (0.32)	-0.68* (0.36)	-0.69* (0.36)	-0.66* (0.36)
Oil exporter <sub>c</sub>	0.64** (0.31)	0.66** (0.31)	0.61** (0.31)	0.73** (0.33)	0.65* (0.33)	0.51 (0.32)
Anocracy <sub>a, c</sub>	0.91*** (0.23)	0.94*** (0.23)	0.97*** (0.23)	0.64** (0.26)	0.62** (0.26)	0.62** (0.26)
Brevity of peace	0.40 (0.26)	0.44* (0.26)	0.51** (0.26)	0.42 (0.29)	0.42 (0.29)	0.46 (0.29)
Constant	-6.42*** (0.72)	-5.96*** (0.75)	-5.68*** (0.71)	-7.22*** (0.79)	-6.91*** (0.81)	-6.72*** (0.78)
Pseudo R <sup>2</sup>	0.082	0.086	0.083	0.089	0.091	0.089
Log likelihood	-393.69	-392.05	-393.41	-325.75	-325.02	-325.81
Observations	5149	5149	5149	4561	4561	4561
Onset Events	84	84	84	69	69	69

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous

It is difficult to assess why our results differ. We both use the Fearon-Laitin measure of civil war onset. A possible, but not very plausible reason is that he includes a smaller number of

<sup>78</sup> He also uses the Fearon-Laitin civil war measure.

observations (at most 4830) than I do (5149). The main difference concerning data is the *per capita income* measure. He only reports that he gets it from PWT 5.6 and imputes missing data from WDI (Kocher 2004:64). Consequently, we can not tell which exact GDP *per capita* estimate he uses out of several available options.

The non-significance of % *urban* is probably much due to its high correlation with per capita income ( $r = 0.85$ ). I exclude *per capita income* from the equation in Model 8-3 to better assess this. The effect of % *urban* then becomes significant at the 0.01 level. This lower p-value is due to its coefficient being almost doubled (presumably because we are no longer comparing countries at the same level of *per capita income*) and due to a sharp decrease of its standard error (because the collinearity problem is gone). Thus, the first part of the hypothesis should also be considered strengthened by the results. The low level of significance of the % *urban* coefficient in Model 8-2 is probably much due to collinearity making the estimate of its individual effect imprecise. *H6* is thus strongly supported by these results.

#### **6.4.2 The telecommunication infrastructure hypothesis (H7)**

In chapter three I discussed an ecological explanation holding that high cost and low availability of information make poor countries particularly prone to civil war (Section 3.3.2). I derived a hypothesis (*H7*) which holds (a) that *lower stocks of telecommunications infrastructure are related to a higher risk of civil war onset*, and (b) that *the per capita income effect on risk of civil war onset is considerably weakened by controlling for telecommunication infrastructure*. I use the number of telephone sets per 1000 inhabitants (“*telephones*”) as a proxy for telecommunication infrastructure (see Section 5.5.2). This variable has to my knowledge not been tested in any quantitative study before.

Model 8-4 above shows the results of regressing civil war onset on the basic model, excluding the observations missing from *telephones*. In Model 8-5, the *telephones* variable is included. Its coefficient has the predicted negative sign, but is not significant even at  $\alpha = 0.10$ . Notably, the *per capita income* coefficient is almost halved and no longer significant ( $p > 0.10$ ). The results show clear signs of high collinearity. First, the standard error of the *per capita income* coefficient has doubled from the prior model. Second, the *telephones* and *per capita income* are correlated by  $r = 0.89$ . Their VIF-values are respectively 5.76 and 6.05 (Appendix 2.5). This is below the common rule of thumb of  $VIF=10$ , but in my model with relatively few onset observations it may be problematic. The two variables are not likely to have very large

independent effects on the dependent variable, and if they do, these will be quite imprecisely estimated in the logit analysis. High collinearity thus probably explains much of the non-significance of their individual coefficients in Model 8-5.

Model 8-6 shows the results of excluding *per capita income* from the prior model. As expected, *telephones* becomes highly significant ( $p < 0.01$ ). This is due to its coefficient being doubled and its standard error being halved. Interestingly, the log likelihood is just barely higher than that of Model 8-4, where only *per capita income* is included. Thus, the effect of the two variables appears to be equally strong, and account for approximately the same amount of variation in the dependent variable.

I do a likelihood ratio test to better assess whether *telephones* has an effect which is independent of the *per capita income* effect. Comparing the -2 log likelihood value (-2LL) of Model 8-5, which includes both variables, and -2LL of Model 8-4, which excludes *telephones* gives the likelihood ratio:

$$LR = (-2LL_{\text{Model8-4}}) - (-2LL_{\text{Model8-5}}) = 1.47.$$

The likelihood ratio has a chi-squared distribution with 1 degree of freedom (df) (the number of parameters added to Model 8-5) (Skog 2004:413). This gives a p-value of 0.23. Hence, *telephones* does not add a significant, independent effect to the model including *per capita income*. Similarly, a likelihood ratio test comparing -2LL of Model 8-5 with -2LL of Model 8-6 shows that *per capita income* does not have a significant, independent effect either.<sup>79</sup>

The results strengthen *H7*. Telecommunication infrastructure is strongly and positively related to *per capita income*, and controlling for the number of telephones per 1000 inhabitants, *per capita income* has no significant effect. Thus, it is possible that poor countries are more prone to civil war due to their low supply of information, which makes guerilla warfare more feasible. It must be noted, however, that *telephones* is a broad measure which probably reflects not only the flow of information, but several other factors relevant to the risk of civil war.

---

<sup>79</sup>  $LR = (-2LL_{\text{Model8-6}}) - (-2LL_{\text{Model8-5}}) = 1.58$ . This is not significant at the 0.10 level ( $p = 0.21$ ).



### 6.4.3 Summary of results for the ecological hypotheses

The prediction concerning *road density* (H4.2) was derived from both the state weakness and the ecological explanation. The test in Section 6.3.2 gave partial support for the hypothesis. The *per capita income* effect was considerably reduced when controlling for *road density*, but *road density* did not have a significant effect when holding *per capita income* at a constant level. The other two hypotheses were strongly supported. Both settlement patterns and telecommunication infrastructure could account for a large part of the *per capita income* effect. This may not be very surprising, as both the proportion of people living in urban areas and telecommunication infrastructure are strongly related to economic development, like *per capita income*. It is therefore difficult to draw strong conclusions regarding the ecological explanation on the basis of these findings. However, these two ecological accounts must be considered strengthened.

## 6.5 Testing the greed hypotheses

The greed explanation holds that poor countries are more prone to civil war primarily because the opportunity cost of rebelling is lower for poor people. Lower economic opportunities in the regular labor market should be associated with a greater (greed) motive and a larger opportunity for rebellion according to this theory. The greed explanation thus holds that poverty (proxied by *per capita income*) should have a direct negative effect on the risk of civil war. I will try to test this hypothesis (H8.1) although it is a difficult task. The greed explanation can be more satisfactorily assessed by also testing other empirical implications of the argument. I have derived predictions concerning *economic growth* (H8.2), *male secondary schooling* (H8.3), *income inequality* (H8.4), and *extraction level* (H8.5) (Section 3.4.3). The *extraction level* hypothesis was tested above (Section 6.3.5). This section provides tests of the remaining hypotheses.

### 6.5.1 The economic growth hypothesis (H8.2)

*Economic growth* may, like *per capita income*, affect the economic opportunities available in the regular labor market. Thus, if the greed account of the poverty-civil war relationship is correct, we should expect economic growth to be positively related to the risk of civil war as well. H8.2 simply holds that *economic growth is negatively related to the risk of civil war onset*.

The quantitative literature is in disagreement about the relationship between economic growth and risk of civil war onset. Collier and Hoeffler (2004:573, 578) find a negative effect of economic growth which is highly significant in most of their models. Their finding is supported by Hegre and Sambanis (2006). In their sensitivity analysis of 88 variables used in the quantitative civil war literature, economic growth comes out as one of the most robust variables (Hegre and Sambanis 2006:33). Sambanis (2004b), on the other hand, does not find growth to be significant. He also reports that its coefficient has a positive sign in half of his regression models (Sambanis 2004b:840).

The endogeneity problem related to this variable is considerable. High prospects of war may start processes that decrease economic growth. People are likely to save a larger proportion of their income to prepare for the crisis and capital may leave the country. Miguel et al. (2004) take an instrumental variable approach to avoid the endogeneity problem. They use variations in rainfall as an instrumental variable for economic growth in 41 African countries in the period 1981-99. Rainfall variation is clearly not endogenous to civil war, as looming conflict does not affect the weather. In the highly agricultural countries of their analysis, rainfall variation is probably a good indicator of economic growth. They find that economic growth has a strong negative impact on the risk of civil war (Miguel et al. 2004:746-747).

This hypothesis can be tested simply by adding *economic growth* to the basic model. Model 9-2 shows the results of this logistic regression. The *economic growth* coefficient has the predicted negative sign, but is not significant at  $\alpha = 0.10$ , though it is very close ( $p = 0.11$ ).<sup>80</sup> It should be noted that this is one of the models where “rare events logit” estimation gives a very different result (see Section 4.4). With this estimation technique the negative economic growth coefficient becomes significant at  $\alpha = 0.05$ . It is difficult to assess why rare events logit produces different estimates particularly here. I can not determine which estimation we should trust, and I will therefore not make a strong inference.

Although I lag the economic growth variable one year, it is far from certain that I avoid endogeneity. Endogeneity would here imply that the negative relationship found between economic growth and civil war onset is partly due to civil war impeding economic growth.

---

<sup>80</sup> Its low significance is not plausibly due to a high degree of collinearity. A likelihood ratio test results in approximately the same p-value (0.12).

Hence, the estimate would be biased toward a *larger* negative causal effect of economic growth on the risk of civil war. This makes it even less certain that the hypothesis is correct.

**Table 9.** Test of H8.2, H8.3, and H8.4: Estimating the effect of economic growth, male secondary schooling, and income inequality on the risk of civil war onset.

Variable	Model 9-1	Model 9-2	Model 9-3	Model 9-4	Model 9-5
Economic growth <sub>a</sub>	-0.027 (0.017)				
Male secondary schooling <sub>a</sub>			-0.021*** (0.0058)	-0.011 (0.0088)	
Income inequality (gini)					-0.0076 (0.017)
Income per capita <sub>a, b</sub>	-0.70*** (0.14)	-0.62*** (0.16)		-0.38 (0.25)	-0.48** (0.19)
Population <sub>b</sub>	0.30*** (0.081)	0.27*** (0.095)	0.33*** (0.096)	0.30*** (0.099)	0.36*** (0.12)
Ongoing war <sub>a</sub>	-0.57* (0.33)	-0.77** (0.38)	-0.74* (0.38)	-0.77** (0.38)	-0.32 (0.42)
Oil exporter <sub>c</sub>	0.38 (0.37)	0.17 (0.41)	-0.033 (0.40)	0.13 (0.41)	0.54 (0.42)
Anocracy <sub>a, c</sub>	0.57** (0.25)	1.084*** (0.29)	1.06*** (0.29)	1.08*** (0.29)	1.20*** (0.33)
Brevity of peace	0.53* (0.27)	0.67** (0.31)	0.80** (0.31)	0.71** (0.31)	0.59 (0.37)
Constant	-6.94*** (0.78)	-6.64*** (0.90)	-6.70*** (0.89)	-6.64*** (0.90)	-7.75*** (1.58)
Pseudo R <sup>2</sup>	0.087	0.089	0.088	0.092	0.10
Log likelihood	-342.03	-249.31	-249.68	-248.51	-201.76
Observations	4840	3183	3183	3183	3269
Onset Events	72	54	54	54	42

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous

### 6.5.2 The male secondary schooling hypothesis (H8.3)

*Male secondary schooling* can also indicate economic opportunities, particularly for young men. As young men often make out a large proportion of soldiers in rebel groups, this should be a highly relevant proxy. H4.3 holds that *male secondary schooling is negatively related to the risk of civil war onset*. I measure *male secondary schooling* by gross male secondary school enrollment rates (Section 5.6.2). Collier and Hoeffler (2004) have previously found a significant negative relationship between *male secondary schooling* and the risk of civil war onset. They even report that substituting *per capita income* with *male secondary schooling* gives the model a slightly better fit to the data (Collier and Hoeffler 2004:573-574).

I first investigate whether *male secondary schooling* may capture some different aspect of economic opportunity than *per capita income*. A bivariate correlation analysis shows that they are highly correlated ( $r = 0.79$ ). *Male secondary schooling* may thus not be expected to have a large effect which is independent of the *per capita income* effect on the dependent variable.

Including both explanatory variables in one model will probably lead to imprecise estimates of their individual effects and high standard errors. I therefore start with a logistic regression of civil war onset on the basic model, excluding the observations missing from *male secondary schooling*. The results are shown in Model 9-2 above.

In Model 9-3, *per capita income* is excluded and *male secondary schooling* is included. *Male secondary schooling* has a negative and highly significant effect ( $p < 0.01$ ) as predicted by *H4.3*. The model's predictive power, shown by the log likelihood, is somewhat weakened when substituting *per capita income* with *male secondary schooling*, however. In Model 9-4 I include both variables. Neither the *per capita income* nor the *male secondary schooling* coefficient reaches significance even at  $\alpha = 0.10$  in this model. A likelihood ratio test comparing Model 9-4 with either of the two prior models shows that none of the coefficients add a significant effect which the other cannot account for.<sup>81</sup> This is an interesting finding, but one which is not directly related to the hypothesis. I may conclude that *H4.3* is supported by the empirical analysis.

#### **6.4 The income inequality hypothesis (H8.4)**

I have argued that higher levels of income inequality should be related to a larger number of people with low opportunities in the regular labor market. If the greed explanation of the poverty-civil war relationship is correct, we should also expect income inequality to be positively related to the risk of civil war onset (*H8.4*). Income inequality is measured by a gini coefficient (Section 5.6.3).

There is considerable disagreement concerning the empirical relationship between inequality and political conflict in general (Lichbach 1989:1050). However, recent quantitative civil war studies agree that income inequality has no significant effect on the risk of civil war (Collier and Hoeffler 2004:576; Fearon and Laitin 2003:85). This is interpreted as evidence that the grievance theory is weak. I have argued to the contrary that inequality is a more direct and plausible proxy for economic opportunity than for grievance (see Section 3.4.3).

Model 9-5 in the table above shows the results of the logistic regression when adding *income inequality* to the basic model. The income inequality coefficient has a *negative* sign, contrary

to the prediction. The effect is far from significant ( $p = 0.65$ ). There are no indications of collinearity affecting the results, as VIF-values are low (Appendix 2.6). A greater problem is posed by the many missing observations for income inequality. Only 42 civil war onsets are included in the model, which is much lower than for the other tests. This suggests that I should not too readily make inferences. However, the results are very dismissive. I do not think it plausible that bias due to a small sample alone is causing such negative results. Thus, I conclude that the hypothesis is not supported.

#### **6.4 The greed explanation's per capita income hypothesis (H8.1)**

Most of the previous hypotheses have concerned *per capita income* in some way. However, only the opportunity cost or greed theory clearly suggests a direct and unconditional negative effect of *per capita income*. The other explanations hold that the *per capita income* effect is indirect, spurious, or conditional on some other variable. I here investigate the greed hypothesis, *H8.1*, which holds that *income per capita is negatively related to the risk of civil war onset even when controlling for the causal variables of the other explanations*.

This hypothesis is obviously difficult to test. There is a large collinearity problem, particularly because *per capita income* is very highly correlated with two of the ecological causal variables – the proportion of people living in urban areas (*% urban*) and telecommunication infrastructure (*telephones*). Also, there is a considerable problem of missing observations when I include many variables in one regression model. This empirical assessment will thus be somewhat tentative.

In Model 10-1 I include all the causal variables from the other explanations which could account for a considerable part of the *per capita income* effect in the above tests. None of these have significant effects in this model. This is not surprising, as the number of observations is low ( $N = 2538$ ) and standard errors are high due to collinearity. Interestingly, both the *per capita income* and the *telephones* coefficient has changed sign and become positive. These are also the variables with the highest level of multicollinearity. The VIF-value of per capita income is 7.71 and the VIF of telephones is 8.90 (Appendix 2.7). Combined with few observations this makes it very difficult to make any inferences. In Model

---

<sup>81</sup> Comparing -2LL of Model 9-4 with -2LL of Model 9-2, which excludes *males secondary schooling*, gives a likelihood ratio of 1.60 with a p-value of 0.21. Comparing Model 9-4 with Model 9-3, which excludes *per capita income*, the likelihood ratio is 2.35, with a p-value of 0.13.

10-2 I exclude *primary schooling*, the variable with the most missing observations. This makes both N and the number of onsets considerably greater. *Income per capita* and *telephones* still have positive coefficients. Notably, the negative *% urban* coefficient becomes significant at the 0.05 level. Hence, *% urban* has a negative effect even when holding *per capita income* and *telephones* constant.

**Table 10.** Test of H8.1: Estimating the effect of per capita income on the risk of civil war onset.

Variable	Model 10-1	Model 10-2	Model 10-3	Model 10-4
Income per capita <sub>a, b</sub>	0.53 (0.48)	0.25 (0.40)	-0.035 (0.35)	0.14 (0.29)
Primary schooling <sub>a</sub>	-0.010 (0.0076)			
Population density <sub>b</sub>	0.25 (0.18)	0.14 (0.15)		0.12 (0.13)
Road density	-2.35 (1.58)	-0.81 (0.82)		-0.81 (0.71)
% Urban	-0.032 (0.020)	-0.034** (0.017)	-0.028* (0.015)	-0.029** (0.012)
Telephones <sub>a, b</sub>	0.039 (0.28)	0.029 (0.21)	-0.033 (0.19)	
Population <sub>b</sub>	0.34*** (0.12)	0.30*** (0.10)	0.29*** (0.085)	0.30*** (0.095)
Ongoing war <sub>a</sub>	-0.60 (0.45)	-0.60 (0.42)	-0.68* (0.38)	-0.68* (0.37)
Oil exporter <sub>c</sub>	-0.14 (0.57)	0.29 (0.45)	0.49 (0.39)	0.42 (0.37)
Anocracy <sub>a, c</sub>	0.57 (0.37)	0.51 (0.31)	0.70** (0.28)	0.82*** (0.27)
Brevity of peace	0.83** (0.38)	0.66* (0.34)	0.53* (0.31)	0.45 (0.31)
Constant	-4.95*** (1.41)	-5.67*** (1.25)	-6.06*** (0.84)	-5.74*** (1.13)
Pseudo R <sup>2</sup>	0.11	0.082	0.079	0.086
Log likelihood	-175.38	-240.07	-285.47	-299.32
Observations	2538	3456	3888	4227
Onset Events	38	50	60	63

Standard errors are in parentheses. \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\* $p < 0.01$  (two-tailed tests).

a = Lagged one year, b = Values are logged, c = Dichotomous

Model 10-3 shows the results of an analysis where *population density* and *road density* have been excluded. Ten additional onset events are then considered. The *per capita income* coefficient becomes negative again, but is very small and far from significant ( $p = 0.63$ ). The same is the case for the *telephones* coefficient. *% Urban* still has a negative effect, but is now only significant at the 0.10 level. It is possible that the higher significance of *% urban* than the other two variables is partly due its lower level of multicollinearity. VIF is 4.45 for *% urban*, 6.88 for *per capita income*, and 7.04 for *telephones* (Appendix 2.9). Finally, in Model 10-4 *telephones* is excluded and *road density* and *population density* are again taken into the analysis. This further increases the number of observations and onsets somewhat. The *income*

*per capita* coefficient turns up as positive, whereas % *urban* has a negative effect which is significant at the 0.05 level. Thus, *per capita income* has no negative effect on the risk of civil war when holding *road density*, *population density* and % *urban* constant.

It is difficult to assess how collinearity is affecting these results. We should probably not make strong inferences concerning the positive effect of *per capita income* in Model 10-2 and Model 10-2. In those models VIF is high and observations are few. However, the results of Model 10-4 is particularly disconfirming for hypothesis 8.1. In this model the number of observations is relatively high, and the VIF-value of *per capita income* is only 4.01. This is barely higher than the VIF of % *urban*, which has a significant effect (Appendix 2.10). The results thus clearly suggest that % *urban* is the more robust variable of the two. Though there is some degree of uncertainty involved, I may conclude that the evidence is not supportive of the greed hypothesis.

### **6.9 Summary of results for the greed hypotheses**

The empirical results were not very supportive of the greed hypotheses. Only the male secondary schooling hypothesis (*H8.3*) received some degree of support. *Male secondary schooling* was negatively related to civil war onset, as predicted. It did not have an effect independent of per capita income, however. I found that *economic growth* (*H8.2*) had the predicted negative effect on the risk of civil war, but was not significant. The effect of *income inequality* had the opposite of the predicted sign, clearly contradicting hypothesis 8.4. The results for the *extraction level* hypothesis (*H8.5*), tested in Section 6.3.5, were also dismissive. Finally, I found that *per capita income* was not negatively related to the risk of civil war onset when controlling for the most significant causal variables of the other explanations. The first greed hypothesis, *H8.1*, thus neither received any empirical support.

## 7 Conclusion

This thesis has explored a grand question: Why do most civil wars take place in poor countries? Scholars have suggested widely different answers to this. I have evaluated four categories of explanations which are prominent in the quantitative civil war literature. First, the *grievance explanation* holds that people in poor countries for various reasons are likely to be more frustrated with their government, which increases their motivation to rebel. Second, the *state weakness explanation* holds that poor countries are more prone to civil war because their state organizations are less capable of controlling social relationships in general and defeating rising insurgencies in particular. Third, what I have termed the *ecological explanation* holds that civil wars more often occur in poor countries because they have social and physical environments more conducive to guerilla warfare, which has been the standard military tactic of 20<sup>th</sup> century insurgents. Finally, the *greed explanation* holds that poor countries have a higher risk of civil war because few economic opportunities imply a greater incentive for people to become rebels, and this in turn makes rebellion more viable.

### 7.1 Summary of findings

My study is the first to systematically assess the relative strength of these explanations. In the theoretical analysis I investigated their assumptions and logical coherence, finding that particularly the greed explanation has a weak theoretical foundation. I also specified which lower-level arguments belonging to the explanations are plausible and consistent. A more precise causal structure was then derived from these arguments. Importantly, I assessed whether the arguments considered the per capita income effect to be direct, indirect, spurious, or conditional. This allowed me to formulate some empirically testable hypotheses for each explanation. The hypotheses concerned not only the relationship between the explanatory variable in question (e.g. state military capacity) and civil war onset. They also specified how the per capita income effect is altered when including this explanatory variable in a statistical analysis. I tested the hypotheses in succession using logistic regression analysis of quantitative data for 161 countries from 1945 to 1999.

In my theoretical analysis of the *grievance explanation*, I argued that the simplest grievance account - that poverty per se makes people more likely to rebel - does not accord with relative deprivation theory. I derived two more theoretically valid grievance accounts. First, *increasing levels of taxation* may create grievances particularly in poor societies because it



conflicts with prevailing norms. Hence, civil wars in poor countries may be explained by people's rage over increased taxation by an expansive state – a modern parallel to the “tax rebellions” in Western Europe from the 16<sup>th</sup> to the 19<sup>th</sup> century. The empirical results quite strongly supported the contention that an increasing level of tax extraction is related to a higher risk of civil war onset. However, the relationship was not any stronger in poorer countries or countries with initially lower levels of extraction, as the hypothesis suggested. Thus, increasing taxation does not appear to account for the poverty-civil war relationship.

The second grievance account holds that a *lack of public goods provision* by the state in poor countries may create grievances that motivate people to rebel. I used primary school enrollment as a proxy for public goods provision. The empirical analysis showed that a higher level of primary school enrollment was significantly related to a lower risk of civil war. Also, the effect of per capita income on the risk of civil war onset was considerably reduced by controlling for primary school enrollment. This suggests that public goods provision may explain some of the poverty-civil war relationship. However, this grievance account rests on a dubious assumption. It is not certain that most people in poor societies expect the government to provide public goods such as primary education. If they do not, the lack of education opportunities probably does not cause intense grievances. I have therefore been reluctant to make inferences, and conclude that this grievance argument can not be dismissed.

The *state weakness explanation* is well founded in the theoretical literature on insurgency and civil war. The argument that poor countries are more prone to civil war because they have a weak state apparatus is theoretically sound. Insurgency should hardly be viable if the state has a strong administrative and military presence throughout its territory. Testing the argument is difficult, however. We lack highly valid and reliable measures of state capacities with good data coverage. I tried to compensate for this by finding several different indicators and testing them separately. Military personnel per capita was used as a proxy for counterinsurgent capacity, road density for the state's reach into rural society, government share of consumption for the state's available resources, and tax-to-GDP ratio for state administrative capacity. I also derived a somewhat different state weakness hypothesis. This holds that *decreasing revenues in weak states* can explain parts of the poverty-civil war relationship.

The state weakness hypotheses received very little support. The per capita income effect was considerably reduced only when controlling for road density. Road density did not have a

significant effect on the risk of civil war, however, holding per capita income constant. Hence, the hypothesis was at best partly supported. Interestingly, a higher number of military personnel per capita *was* related to a lower likelihood of civil war. This was largely independent of the per capita income effect, however. Government share of consumption and tax ratio did not have a significant effect on the risk of civil war. The results were also clearly dismissive of the hypothesis concerning decreasing revenues in weak states. These negative findings should be prudently interpreted, as the validity of all these measures can be questioned. Still, if state weakness could truly account for a large part of the poverty-civil war relationship, we should expect to find empirical support for at least some of the hypotheses. The state weakness explanation must therefore be considered weakened by the results.

I find that the *ecological explanation* rests upon a theoretical framework quite similar to the state weakness explanation. They both emphasize the politico-military aspects of civil war, and understand insurgency primarily as a form of competitive state building. I derived three different ecological accounts. The first of these gives the same prediction as the state weakness explanation: a lower *road density* in poor countries may explain their higher propensity to experience civil war. In an ecological perspective, low road density should be conducive to insurgency because it increases the possibility for guerillas to hide from government forces in the periphery. This hypothesis was only partly supported, as discussed above. The second ecological account holds that poor countries are more prone to civil war because they tend to have a rural, village-based *settlement pattern*, which is more favorable to insurgency than an urban pattern settlement. This hypothesis was strongly supported in the statistical analysis. The per capita income effect was almost halved when controlling for the proportion of people living in urban areas. A higher proportion of people in urban areas was also significantly related to a lower risk of civil war despite a quite high level of collinearity making the estimates imprecise.

The second ecological hypothesis concerned *telecommunication infrastructure*. I argued that an environment characterized by low availability and high cost of information is conducive to insurgents using guerilla warfare. Accordingly, poor countries may face a higher risk of civil war because they have less extensive telecommunication infrastructure. This hypothesis was also highly supported. The per capita income effect was almost halved when holding constant the number of telephones per thousand inhabitants. As predicted, more telecommunication infrastructure was related to a lower risk of civil war onset. Its effect was not significant when

including per capita income, but this was much due to multicollinearity making estimates uncertain. Hence, the results can be cited in support of the argument that availability of information explains parts of the poverty-civil war puzzle. Two of the ecological hypotheses were thus clearly strengthened by the empirical findings.

In the theoretical analysis I argued that the *greed explanation* is not readily derivable from economic conflict theory. It is not certain that lower economic opportunities in the regular labor market will generally favor the rebels vis-à-vis the state. A low-income economy implies lower labor costs not only for a potential rebel organization, but also for the state. In consequence, the greed explanation rests on Collier's (2000b) uncertain assumption that rebel organizations are more sensitive to labor costs than is the government. Though theoretically fragile, this account can not be a priori dismissed. I therefore tested several empirical implications of the economic opportunity argument. Five indicators of economic opportunity were suggested: Male secondary schooling, economic growth, income inequality, level of taxation, and finally, per capita income. To the extent that per capita income reflects economic opportunity, the greed explanation predicts a direct and unconditional causal effect of per capita income on the risk of civil war onset. The other economic opportunity variables were not hypothesized to weaken or condition the per capita income effect; they are different implications of the same argument, and therefore useful for assessing the greed explanation.

The *male secondary schooling* hypothesis received the most empirical support of the greed hypotheses. A higher rate of male secondary school enrollment was related to a lower risk of civil war. However, this variable did not have any effect independent of per capita income. *Economic growth* was also negatively related to the risk of civil war onset as predicted, but the effect was not significant. The third economic opportunity indicator, *income inequality*, is more controversial than the others. The foremost greed theory proponents, Collier and Hoeffler, use income inequality as a proxy for grievances, not greed. I argue to the contrary that it is more clearly related to economic opportunity. Holding per capita income constant, a higher level of income inequality necessarily implies a larger number of people with a low income. Income inequality may in fact be the best economic opportunity indicator available. Unlike per capita income, male secondary schooling and economic growth, it is not plausibly related to state capacity. It can therefore help us distinguish between these two theories. The results clearly contradicted the income inequality hypothesis. There was even a non-

significant tendency that a higher level of income inequality was related to a *lower* risk of civil war onset.

The fourth economic opportunity indicator I suggested was *level of taxation*, which I test using the tax-to-GDP ratio. I argued that a higher level of taxation implies a lower opportunity cost of rebelling, as rebels do not pay taxes. The economic opportunity account should therefore predict that taxation is *positively* related to the risk of civil war onset, contrary to the prediction of the state weakness explanation. This argument may be controversial, as proponents of greed theory hold that increased state capacity should to some extent also reduce the opportunity for rebelling. I therefore controlled for state military capacity when testing the taxation hypothesis. Tax ratio was still negatively, but not significantly, related to the risk of civil war. Hence, this greed hypothesis was clearly not supported.

The final greed hypothesis concerned *per capita income*. As the greed explanation predicts the opposite of the other explanations concerning this variable, this allows for something reminiscent of a critical test. If per capita income has a direct and unconditional causal effect, as the greed account predicts, it should be negatively related to the risk of civil war even when controlling for the causal variables of the other explanations. I included the four variables found to be most relevant from the previous analyses. Holding these explanatory variables constant, the per capita income effect (in three of the models) changed sign and became positive, or (in one model) approached zero.

There are clearly caveats to this final test. Three of the variables (% urban, telephones, and per capita income) were quite highly correlated. Hence, it was not very likely that they had large independent effects. The high degree of multicollinearity also made estimates less precise. Still, the proportion of people living in urban areas had a negative *and significant* effect in all the three models with the most observations. Hence, settlement pattern seems to be the most robust of the explanatory variables. This result clearly strengthens the ecological explanation.

## 7.2 Discussion

There are two main problems that limit the feasible range of inferences from the empirical analysis. First, we lack measures that we know are highly valid and reliable for some of the theoretical concepts. This is a *data problem* common to most cross-national research. The shortcomings of state capacity indicators are perhaps the most serious here. Clearly, per capita income is a very crude indicator of state capacity. I found some previously untested indicators that may be more specific and valid indicators of this concept. However, these have the opposite problem of the per capita income measure; they may be too narrow, thus not capturing the entire state capacity concept. In addition, their reliability is probably not very high. The state capacity explanation should therefore not be promptly dismissed on the basis of the negative statistical results. We need to find new and better operationalizations of state capacity and increase the reliability and data coverage of the available indicators.

The second problem is the high correlation between indicators of the different theoretical variables. This is not a data problem, but derives from social reality. For example, as there is a tendency that wealthier societies are also more urbanized, it is difficult to assess whether the first or the latter characteristic is causing a lower risk of civil war in these societies. However, I found that no variables were completely overlapping. Urban settlement had a significant effect even when controlling for per capita income and telecommunication infrastructure. I also find a surprisingly low correlation between per capita income and the state capacity indicators. Hence, in many of the analyses the problem is insignificant.

These two problems, lack of valid measures and hardly distinguishable explanations, were the main challenges to my study. With a different research design, I might have been able to handle them more satisfactorily. Taking the difficulty of separating the causal effect of urban settlement and poverty as an example, a historical comparative case study design could have been useful. It would be interesting to compare a few countries that over some decades go from being poor and rural to being poor and urbanized, with countries that go from being poor and rural to being both more wealthy and urbanized. Such a design would be focused on cases where the explanatory variables take different values, which helps us separate the explanations. Also, in-depth studies may enable us to use more fine-grained and valid measures, thus reducing the first problem.

This being said, my research design has clear advantages compared with such an alternative design. First, the usual objections to small-N designs apply also here. They reduce the possibility to control for many variables and they involve a risk of selection bias. More importantly perhaps, a case study could hardly have taken on the grand task of comparing four categories of explanations for the poverty-civil war relationship. I found my large-N approach most suitable for this. One might object that the task was overambitious. However, my research question and design was chosen due to the shortcoming of previous studies; they had largely promoted their one great explanation without systematically comparing it to rival explanations. This left us bewildered. We could not know which of several different explanations of the poverty-civil war relationship was the strongest. The importance of the task may therefore to some extent justify my choice.

More importantly, I have not found the problems facing my study insurmountable. I would argue that it has brought us one step further in unveiling the poverty-civil war relationship. My study has shown that the two most prominent explanations in the quantitative literature, state weakness and greed, are probably not the strongest. The state weakness explanation is theoretically sound, but found very little empirical support. The greed explanation is not theoretically convincing, and the empirical results were also largely in its disfavor. Concerning the grievance explanation, only the grievance argument focusing on public goods provision received some support. This argument rests on a dubious assumption, however, which reduces its credibility.

My analysis suggests that the ecological explanation of the poverty-civil war relationship is the strongest of the four. It is highly supported by the empirical evidence; holding the proportion of people living in urban areas and the stocks of telecommunication infrastructure constant, the effect of per capita income disappears. I also find the explanation theoretically convincing; an environment characterized by high availability of information is not favorable to insurgents because they are, at least initially, dependent on hiding from government forces. Civil war thus appears to be “a problem of the poor” mainly because low-income countries have social and physical environments conducive to insurgency and guerilla warfare. This suggests that civil war may become a less common manifestation of conflict in the 21<sup>st</sup> century. As more countries become highly urbanized and information technology spreads, political dissidents and opportunists may find other and hopefully less destructive strategies more expedient than insurgency to achieve their goals.

# References

Arbetman, Marina and Jacek Kugler (eds.) (1997). *Political capacity and economic behavior*. Boulder, Colorado: Westview Press.

Acemoglu, Daron and James A. Robinson (2001). "A theory of political transitions." *The American Economic Review* 91(4): 938-963.

Azam, Jean-Paul (2001). "The redistributive state and conflicts in africa." *Journal of Peace Research* 38(4): 429-444.

Azam, Jean-Paul (2006). "The paradox of power reconsidered: a theory of political regimes in Africa." *Journal of African Economies* 15(1): 26-58.

Azam, Jean-Paul and Anke Hoeffler (2002). "Violence against civilians in civil wars: looting or terror?" *Journal of Peace Research* 39(4): 461-485.

Azam, Jean-Paul and Alice Mesnard (2003). "Civil war and the social contract", *Public Choice* 115: 455-475.

Ayoob, Mohammed (1996). "State making, state breaking, and state failure", pp. 37-51 in *Managing global chaos. Sources of and responses to international conflict*. C.A. Crocker, F.O. Hampson, and P.Aall (eds.). Washington, DC: United States Institute of Peace.

Bates, Robert H. (2001). *Prosperity and Violence: The Political Economy of Development*. New York: W.W. Norton.

Buhaug, Halvard and Scott Gates (2002). "The geography of civil war", *Journal of Peace Research* 39(4): 417-433.

Buhaug, Halvard and Jan Ketil Rød (2006). "Local determinants of African civil wars, 1970-2001." *Political Geography* 25: 315-335.

Canning, David (1998). "A Database of World Stocks of Infrastructure, 1950-95." *The World Bank Economic Review* 12(3): 529-47.

Carnoy, Martin (1984). *The state and political theory*. Princeton: Princeton University Press.

Chabal, Patrick and Jean-Pascal Daloz (1999). *Africa works. Disorder as political instrument*. Oxford: James Currey.

Chabal, Patrick and Jean-Pascal Daloz (2006). *Culture troubles. Politics and the interpretation of Meaning*. London: Hurst & Company.

Chaudhry, Kiren Aziz (1997). *The price of wealth: economies and institutions in the Middle East*. Ithaca: Cornell University Press.

Cohen, Youssef, Brian R. Brown, and A.F.K. Organski (1981). "The paradoxical nature of state making: the violent creation of order." *American Political Science Review* 75(4): 901-910.

Coleman, James S. (1990). *Foundations of social theory*. Cambridge, Massachusetts: Harvard University Press.

Collier, Paul (2000a). "Doing well out of war: an economic perspective." In: *Greed and grievance: economic agendas in civil wars*. Mats Berdal og David D. Malone (Eds.). Boulder: Lynne Rienner.

Collier, Paul (2000b). "Rebellion as a quasi-criminal activity". *Journal of Conflict Resolution*, 44(6): 839-853.

Collier, Paul and Anke Hoeffler (2004). "Greed and grievance in civil war." *Oxford Economic Papers*, 56: 563-595.

Collier, Paul and Anke Hoeffler (1998). "On economic causes of civil war." *Oxford Economic Papers*, 50: 563-573.

Collier, Paul and Anke Hoeffler (2007). "Unintended consequences: does aid promote arms races?" *Oxford bulletin of economics and statistics* 69(1):1-27.

Correlates of war (2005). "National military capabilities data documentation. Version 3.0." <http://www.correlatesofwar.org> (27.08.2007).

Davies, James C. (1962): "Toward a theory of revolution." *American sociological review* 27(1): 5-19.

Damgaard, Christian (2003). "Gini coefficient". *MathWorld - A Wolfram Web Resource*. <http://mathworld.wolfram.com/GiniCoefficient.html> (29.09.2007).

Esty, Daniel C., Jack A. Goldstone, Ted R. Gurr, Pamela T. Surko and Alan N. Unger (1995). "State failure task force report." McLean, Science Applications International Corporation, 30 November.

Englebert, P. (2000). *State legitimacy and development in Africa*. Boulder, CO: Lynne Rienner Publishers.

Evans, P. B., D. Rueschemeyer, and T. Skocpol (eds.) (1985): *Bringing the state back in*. Cambridge: Cambridge University Press.

Fauvelle-Aymar, C. (1999). "The political and tax capacity of government in developing countries." *Kyklos* 52: 391-413.

Fearon, James (1995). "Rationalist explanations for war." *International Organization* 49(3): 379-414.



Fearon, James D. and David Laitin (1999). "Weak states, rough terrain, and large-scale ethnic violence since 1945." Paper presented at the annual meetings of the American Political Science Association, Atlanta, GA.

Fearon, James D and David Laitin (2003a). "Ethnicity, insurgency, and civil war." *American Political Science Review* 97(1): 75-90.

Fearon, James D. and David Laitin (2003b). "Additional tables for "Ethnicity, insurgency, and civil war."" <http://www.stanford.edu/group/ethnic/workingpapers/papers.htm> (18.03.2007).

Garfinkel, Michelle R. and Stergios Skaperdas (2006). "Economics of conflict: an overview." In Sandler, T. and Hartley, K. (eds): *Handbook of Defence Economics, Vol. II*. New York: Elsevier.

Gates, Scott (2002). "Empirically assessing the causes of civil war." Paper presented at the annual meeting of the International Studies Association, New Orleans, 24-27 March.

Gleditsch, Kristian S. (2007). "Transnational dimensions of civil war." *Journal of Peace Research* 44(3): 293-309.

Gleditsch, Kristian S. and Michael D. Ward (2001). "Measuring space: a minimum distance database." *Journal of Peace Research* 38(6): 749-68.

Gleditsch, Nils Petter; Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg & Håvard Strand (2002). "Armed conflict 1946–2001: a new dataset." *Journal of Peace Research* 39(5): 615-637.

Greene, William H. (1997): *Econometric analysis*. New York: New York University.

Grossman, Herschel I. (1991). "A general equilibrium model of insurrections." *American Economic Review* 81(4): 912-921.

Gujarati, Damodar N. (2003). *Basic econometrics*. New York: McGraw-Hill/Irwin.

Gurr, Ted R. (1970). *Why men rebel*. New Jersey: Princeton University Press.

Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch (2001). "Toward a democratic civil peace? Democracy, political change, and civil war, 1816-1992." *American Political Science Review* 95(1): 33-48.

Hegre, Håvard and Nicholas Sambanis (2006). "Sensitivity analysis of the empirical literature on civil war onset" *Journal of Conflict Resolution*, 50(4): 508-535.

Henderson, Errol A. and J. David Singer (2000). "Civil war in the post-colonial world, 1946-92." *Journal of Peace Research* 37(3): 275-299.

Herbst, Jeffrey (1990). "War and the state in Africa." *International Security* 14(4): 117-139.

- Herbst, Jeffrey (2000a). "Economic incentives, natural resources and conflict in Africa" *Journal of African Economies* 9(3): 270–294.
- Herbst, Jeffrey (2000b). *States and power in Africa: comparative lessons in authority and control*. Princeton, NJ: Princeton University Press.
- Herbst, Jeffrey (2004). "African militaries and rebellion: the political economy of threat and combat effectiveness." *Journal of Peace Research* 41(3): 357-369.
- Heston et al. (2004). "Data appendix for a space-time system of national accounts: Penn World Table 6.1 (PWT 6.1)". <http://pwt.econ.upenn.edu/Documentation/append61.pdf> (08.09.2007).
- Heston, Alan, Robert Summers and Bettina Aten (2006). *Penn World Table Version 6.2*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. [http://pwt.econ.upenn.edu/php\\_site/pwt\\_index.php](http://pwt.econ.upenn.edu/php_site/pwt_index.php) (08.09.2007).
- Hibbs, Douglas (1973). *Mass political violence*. New York: Wiley.
- Hirshleifer, Jack (2001). *The dark side of the force. Economic foundations of conflict theory*. Cambridge: Cambridge University Press.
- Hobbes, Thomas (1969). *Elements of law, natural and politic*. 2<sup>nd</sup> edition. Ferdinand Tönnies (ed.). Cambridge: Cambridge University Press.
- Holsti, K. (1996). *The state, war, and the state of war*. Cambridge: Cambridge University Press.
- Humphreys, Macartan (2005). "Natural resources, conflict, and conflict resolution: uncovering the mechanisms." *Journal of Conflict Resolution* 49(4): 508-537.
- Huntington, Samuel (1968). *Political Order in Changing Societies*. New Haven: Yale University Press.
- IMF (2001). "Appendix 1: Changes from the 1986 "A Manual on Government Finance Statistics"." <http://www.imf.org/external/pubs/ft/gfs/manual/pdf/app1.pdf> (02.10.2007).
- Jackman, Robert W. (1993). *Power without force: the political capacity of nation-states*. Michigan: University of Michigan.
- Jackson, Robert H. and Carl G. Rosberg (1982). "Why Africa's weak states persist: the empirical and juridical in statehood." *World Politics* 35(1): 1-24.
- Johnson, Kristin and Marina Arbetman Rabinowitz (2005). "Relative Political Capacity Model & Data Description." Paper prepared for the Workshop on Political Economy, Claremont Graduate University, October 7.
- Kalyvas, Stathis N. (2006). *The logic of violence in civil war*. New York: Cambridge University.

- King, Gary, Robert O. Keohane, and Sidney Verba (1994). *Designing social inquiry. Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Kleinbaum, David G., Lawrence L. Kupper, and Keith E. Muller (1988). *Applied regression analysis and other multivariable methods*. Belmont, California: Duxbury Press.
- Krasner, S.D. (1978). *Defending the national interest*. Princeton, NJ: Princeton University Press.
- King, Gary and Langche Zeng (2001a). "Logistic regression in rare events data." *Political Analysis* 9(2): 137-163.
- King, Gary and Langche Zeng (2001b). "Explaining rare events in international relations." *International Organization* 55(3): 693–715.
- Kirby, Andrew and Michael D. Ward (1991). "Modernity and the process of state formation: an examination of 20<sup>th</sup> century Africa." *International Interactions* 17(1) 1991.
- Kocher, Matthew A. (2004). *Human ecology and civil war*. Chicago: University of Chicago, Department of political science. Ph.D. dissertation.
- Lacquer, W. (1977). *Guerrilla: a historical and critical study*. London: Weidenfeld & Nicolson.
- Leites, N. and C. Wolf (1970). *Rebellion and authority: an analytic essay on insurgent conflicts*. Chicago: Markham Publishing.
- Levi, Margareth (1988). *Of rule and revenue*. Berkeley: University of California Press.
- Lichbach, Mark I. (1989). "An evaluation of "does economic inequality breed political conflict?" studies." *World Politics* 41(4): 1049-1076.
- Mehta, Pratap Bhanu (2003). *The burden of democracy*. New Delhi: Penguin Books India.
- Menard, Scott (2004). "Six approaches to calculating standardized logistic regression coefficients." *The American Statistician*. [http://goliath.ecnext.com/coms2/summary\\_0199-762729\\_ITM](http://goliath.ecnext.com/coms2/summary_0199-762729_ITM) (03.10.2007)
- Migdal, J.S. (1988). *Strong societies and weak states: state-society relations and state capabilities in the Third World*. Princeton, NJ: Princeton University Press.
- Miguel, E., S. Satyanath, and E. Sergenti (2004). "Economic shocks and civil conflict: an instrumental variable approach". *Journal of Political Economy* 112(4): 725-753.
- Mousnier, Roland (1970). *Peasant Uprisings in Seventeenth-Century France, Russia, and China*. New York: Harper Torchbooks
- Muller, Edward N., Henry A. Dietz and Steven E. Finkel (1991). "Discontent and the expected utility of rebellion: the case of Peru." *The American Political Science Review*, 85(4): 1261-1282.

Nordhaus, William D. (2007). "Geographically based economic data (G-Econ)".  
<http://gecon.yale.edu/> (13.09.2007).

Olson, Mancur (1965). *The logic of collective action: public goods and the theory of groups*. Cambridge: Harvard University Press.

Organski, A.F.K. and Jacek Kugler (1980). *The War Ledger*. Chicago: Chicago University Press.

Peters, Krijn, Paul Richards, and Koen Vlassenroot (2003). "What happens to youth during and after wars? A preliminary review of literature on Africa and an assessment of the debate." RAWOO Working Paper.  
[http://library.wur.nl/wasp/bestanden/LUWPUBRD\\_00335821\\_A502\\_001.pdf](http://library.wur.nl/wasp/bestanden/LUWPUBRD_00335821_A502_001.pdf). (01.10.2007)

Raknerud, Arvid and Håvard Hegre (1997). "The hazard of war: reassessing the evidence for the democratic peace", *Journal of Peace Research* 34(4): 385-404.

Reno, William (1998). *Warlord politics and African states*. Boulder, CO: Lynne Rienner.

Rokkan, Stein (1975). "Dimensions of state formation and nation-building: A possible paradigm for research on variations within Europe." In Tilly, Charles (ed.): *The formation of national states in Western Europe*. Princeton: Princeton University Press.

Ross, Michael (2004). "What do we know about natural resources and civil war?" *Journal of Peace Research* 41(3): 337-56.

Ross, Michael L. (2006). "A closer look at oil, diamonds, and civil war." *Annual Review of Political Science* 9: 265-300.

Rotberg, Robert I. (ed.) (2004). *When states fail: causes and consequences*. Princeton: Princeton University Press.

Rubinson, Richard (1976). *The world economy and the distribution*

Sambanis, Nicholas (2002). "A review of recent advances and future directions in the quantitative literature on civil wars." *Defence and Peace Economics* 13(3): 215-243.

Sambanis, Nicholas (2004a): "Using case studies to expand the theory of civil war" *Perspectives on Politics* 2(2): 259-279.

Sambanis, Nicholas (2004b). "What is a civil war? Conceptual and empirical complexities of an operational definition." *Journal of Conflict Resolution* 48(6): 814-858.

James C. Scott (1976). *The Moral Economy of the Peasant: Subsistence and Rebellion in Southeast Asia*. New Haven: Yale University Press.

Singer, J. David. (1987). "Reconstructing the Correlates of War Dataset on material capabilities of states, 1816-1985" *International Interactions*, 14: 115-32. Data available at: <http://www.correlatesofwar.org/> (29.08.2007).

Skaperdas, Stergios (1992). "Cooperation, conflict, and power in the absence of property rights." *American Economic Review* 82(4): 720-39.

Skog, Ole-Jørgen (2004): *Å forklare sosiale fenomener: en regresjonsbasert tilnærming*. Oslo: Gyldendal Norsk Forlag.

Small, Melvin, and J. David Singer (1982). *Resort to arms: international and civil wars, 1816-1980*. Beverly Hills, California: Sage Publications.

Spalter, Anne (2007). "Using linear interpolation."  
<http://www.cs.brown.edu/stc/outrea/greenhouse/nursery/interpolation/home.html>.  
(16.10.2007)

Sprout, Harold and Margaret Sprout (1965). *The ecological perspective on human affairs*. Princeton: Princeton university press.

Strand, Håvard (2007). "Retreating from a civil democratic peace? Revisiting the relationship between political institutions and civil war." Draft paper presented at the International Peace Research Institute, Oslo (PRIO) 17.10.2007.

Thies, Cameron G. (2004). "State building, interstate and intrastate rivalry: a study of post-colonial developing country extractive efforts, 1975-2000." *International Studies Quarterly* 48(1): 53-72.

Tilly, Charles (1973). "Do communities act?" IN: *Sociological inquiry* 43(3-4):209-240.

Tilly, Charles (1975). "Reflections on the history of European state making." in Tilly, Charles (ed.): *The formation of national states in Western Europe*. Princeton: Princeton University Press.

Tilly, Charles (1985). "War making and state making as organized crime". In *Bringing the state back In*, P.B. Evans, D. Rueschemeyer and T. Skocpol (eds.), pp. 169-91. Cambridge: Cambridge University Press

Tilly, Charles (1992). *Coercion, capital and European states, AD 990-1992*. Cambridge, MA: Blackwell

Tilly, Charles (2003). *The politics of collective violence*. Cambridge: Cambridge University Press.

Thyne, Clayton (2006). "ABC's, 123's and the golden rule: the pacifying effect of education on civil conflict, 1980—1999." *International Studies Quarterly* 50(4): 733-754.

Tullock, Gordon (1971). "The paradox of revolution." *Public Choice* 11: 89-100.

Urdal, Henrik (2005). "People vs. Malthus: population pressure, environmental degradation, and armed conflict revisited." *Journal of Peace Research* 42(4): 417–434.

United Nations (1989). "Convention on the rights of the child".  
<http://www.ohchr.org/english/law/pdf/crc.pdf> (1.09.2007).

Weber M. (1964). "Politics as a vocation." pp. 77-128 in Gerth, CW Mills (ed.): *From Max Weber: essays in sociology*. New York: Oxford University Press

World Bank (2000). *Entering the 21st Century: World development report 1999/2000*. Washington DC: Oxford University Press.

World Bank (2003). *World development indicators 2003*. Washington, DC: Development Data Center, International Bank for Reconstruction and Development.

World Bank (2006). *World development indicators 2006*. Washington, DC: Development Data Center, International Bank for Reconstruction and Development.

UNESCO (2000). *World education indicators*.  
<http://www.uis.unesco.org/en/stats/statistics/indicators/indic0.htm>. (20.08.2007)

UNU/WIDER (2007). *World income inequality database, version 2.0b*.  
<http://www.wider.unu.edu/wiid/wiid.htm> (19.09.2007).

# Appendices

## 1. Descriptive Statistics

Variables	N	Mean	Standard deviation	Min	Max	Countries included	Onsets included
Income per capita in 1000s	6373	3.65	4.54	0.048	66.74	156	110
Log (Income per capita)	6373	0.74	1.06	-3.04	4.20	156	110
Economic growth	4994	1.91	7.50	-63.32	151.07	153	74
Tax ratio	4206	17.65	9.36	0.4	59	129	66
Tax ratio change (3-1 year before) <small>outliers dropped</small>	3799	-2.30	25.25	-99.55	193.79	125	61
Tax ratio change *	3799	0.25	2.66	-28.39	39.32	125	61
Tax ratio							
Tax ratio change * Log (Income per capita)	3771	-1.46	27.78	-330.05	201.77	121	60
Primary schooling	3558	89.83	27.79	8	200.4	152	60
Male secondary schooling	3277	49.21	30.88	1.7	150.4	145	57
Inequality (Gini)	3322	41.64	11.00	15.9	76.2	137	44
Government share of consumption	4071	15.39	6.55	2.9	76.22	152	67
Decreasing revenues dummy	3612	0.49	0.50	0	1	149	58
*Military personnel per 1000 inhabitants	6513	7.11	9.48	0	210	161	114
Log (Military personnel per 1000 inhabitants)	6344	1.51	1.0010	-2.34	5.35	161	113
Change in military personnel per 1000 inhabitants (3-1 year before)	5849	5.18	35.59	-100	296.03	161	91
Roads density	4512	0.28	0.62	0	4.56	151	65
*Telephones per 1000 inhabitants	4615	85.14	150.32	0	916.95	129	71
Log (Telephones per 1000 inhabitants)	4598	2.97	1.91	-2.14	6.82	129	70
% Urban	5325	44.71	24.17	2.25	1000	157	88
Oil exporter	6610	0.13	0.34	0	1	161	114
Anocracy	6541	0.23	0.42	0	1	161	113
Neighborhood wars	6566	0.58	0.88	0	7	161	114
Ongoing war	6610	0.14	0.34	0	1	161	114
Brevity of peace	6610	84.35	33.44	0	100	161	114
Population density	6587	0.099	0.29	0.00045	5.90	161	114
Log (Population density)	6587	-3.31	1.45	-7.70	1.78	161	114
*Population	6587	31963.3	103345.6	222	1253735	161	114
Log (Population)	6585	9.05	1.46	5.40	14.03	161	114

Operationalizations not used in any analyses are marked by \*.

## 2. Multicollinearity assessment

The VIF-values are estimated from regressing each explanatory variable (X) on all the other explanatory variables (X) in the same model (Section 4.2). See the models in Chapter 6 to find all the variables included.

### 2.1 VIF-values for central explanatory variables in Model 3-4

Variable	VIF ( $1 / 1 - R^2$ )
Tax ratio * tax ratio change	1.17
Tax ratio change	1.15
Tax ratio	1.36
Per capita income	1.18

### 2.2 VIF-values for central explanatory variables in Model 3-5

Variable	VIF ( $1 / 1 - R^2$ )
Tax ratio * Per capita income	1.08
Tax ratio change	1.07
Per capita income	1.15

### 2.3 VIF-values for central explanatory variables in Model 5-4.

Variable	VIF ( $1 / 1 - R^2$ )
Road density	1.97
Population density	1.64
Per capita income	1.54

### 2.4 VIF-values for central explanatory variables in Model 6-2.

Variable	VIF ( $1 / 1 - R^2$ )
Government share	1.15
Per capita income	1.19

### 2.5 VIF-values for central explanatory variables in Model 8-5.

Variable	VIF ( $1 / 1 - R^2$ )
Telephones	5.76
Per capita income	6.05

### 2.6 VIF-values for central explanatory variables in Model 9-5.

Variable	VIF ( $1 / 1 - R^2$ )
Income inequality	1.18
Per capita income	1.26

### 2.7 VIF-values for central explanatory variables in Model 10-1.

Variable	VIF ( $1 / 1 - R^2$ )
Per capita income	7.71
Primary schooling	1.61



Population density	1.67
Road density	1.99
% Urban	4.20
Telephones	8.90

## 2.8 VIF-values for central explanatory variables in Model 10-2.

Variable	VIF ( $1 / 1 - R^2$ )
Per capita income	7.17
Population density	1.53
Road density	1.95
% Urban	4.37
Telephones	7.50

## 2.9 VIF-values for central explanatory variables in Model 10-3.

Variable	VIF ( $1 / 1 - R^2$ )
Per capita income	6.88
% Urban	4.45
Telephones	7.04

## 2.10 VIF-values for central explanatory variables in Model 10-4.

Variable	VIF ( $1 / 1 - R^2$ )
Per capita income	4.01
% Urban	3.66
Road density	2.04
Population density	1.60